(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) **1-0029980**

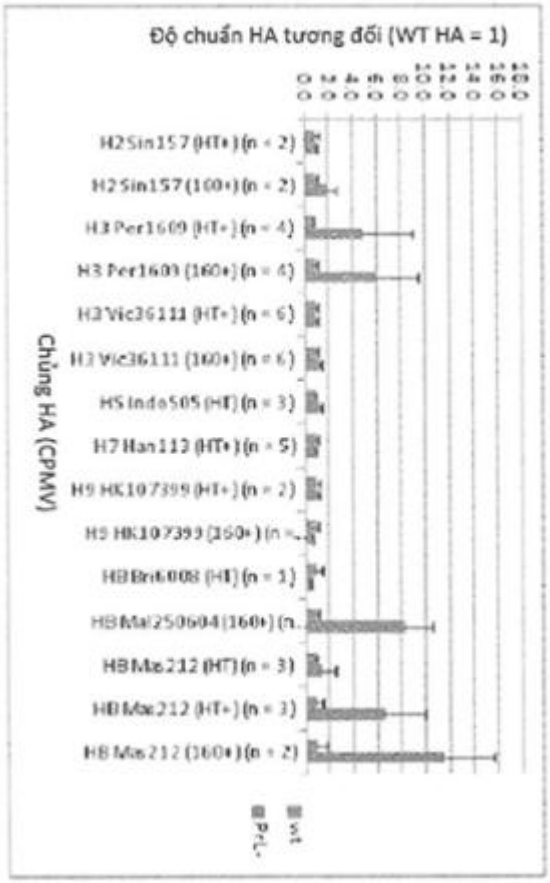
(51)[7] **C12N 15/82**; A61K 39/145; A61P 31/16; A61P 37/04; C12N 7/04; C12N 15/44; C12N 7/01; A01H 5/00; C07K 14/11

(13) **B**

(21) 1-2015-04038
(22) 28/03/2014
(86) PCT/CA2014/050326 28/03/2014
(87) WO2014/153674 02/10/2014
(30) 61/806,227 28/03/2013 US; 61/925,852 10/01/2014 US; 61/971,274 27/03/2014 US
(45) 25/11/2021 404
(43) 25/01/2016 334A
(73) MEDICAGO INC. (CA)
1020 Route de l'Eglise, Suite 600, Quebec, Quebec G1V 3V9, Canada
(72) COUTURE, Manon (CA); D'AOUST, Marc-Andre (CA); VEZINA, Louis-Philippe (CA).
(74) Công ty Luật TNHH T&G (TGVN)

(54) HẠT TƯƠNG TỰ VIRUT (VLP) CÚM Ở THỰC VẬT, PHƯƠNG PHÁP SẢN XUẤT HẠT TƯƠNG TỰ VIRUT, AXIT NUCLEIC ĐƯỢC SỬ DỤNG TRONG SẢN XUẤT HẠT TƯƠNG TỰ VIRUT VÀ VACXIN VÀ CHẾ PHẨM CHỨA HẠT TƯƠNG TỰ VIRUT

(57) Sáng chế đề cập đến phương pháp tạo ra hạt tương tự virut (virus like particle -VLP) ở thực vật bao gồm hemagglutinin được cải biến. Phương pháp này bao gồm việc đưa axit nucleic bao gồm vùng điều hoà có hoạt tính ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá protein hemagglutinin (HA) cúm được cải biến vào trong thực vật, hoặc phần của thực vật, protein HA được cải biến bao gồm vòng phân giải protein được cải biến. Tiếp theo, ủ thực vật hoặc phần của thực vật trong các điều kiện mà cho phép sự biểu hiện của axit nucleic, do đó tạo ra VLP. Sáng chế cũng đề cập đến hạt tương tự virut (VLP) được tạo ra bằng phương pháp này và thực vật biểu hiện VLP này. Sáng chế cũng đề cập đến chế phẩm và vacxin chứa VLP này và axit nucleic được sử dụng trong sản xuất VLP này.


Độ chuẩn HA tương đối (WT HA = 1)
Chủng HA (CPMV)
H2Sin157 (HT+) (n = 2)
H2Sin157 (160+) (n = 2)
H3Per1609 (HT+) (n = 4)
H3Per1609 (160+) (n = 4)
H3Vic36111 (HT+) (n = 6)
H3Vic36111 (160+) (n = 6)
H5Indo505 (HT) (n = 3)
H7Han113 (HT+) (n = 5)
H9HK107399 (HT+) (n = 2)
H9HK107399 (160+) (n =
HBBri6008 (HT) (n = 1)
HBMal250604 (160+) (n
HBMal212 (HT) (n = 3)
HBMal212 (HT+) (n = 3)
HBMal212 (160+) (n = 2)

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp sản xuất hạt tương tự virut ở thực vật.

Tình trạng kỹ thuật của sáng chế

Bệnh cúm do ARN virut thuộc họ orthomyxoviridae gây ra. Có ba typ trong số các virut này và chúng gây ra ba typ cúm khác nhau: typ A, B và C. Các virut typ A của virut cúm lây nhiễm cho động vật có vú (con người, lợn, chồn, ngựa) và chim. Typ này là rất quan trọng đối với loài người, do đây là typ virut mà gây ra các bệnh dịch lớn trên toàn thế giới. Virut cúm typ B (cũng được biết đơn giản là cúm B) chỉ lây nhiễm cho con người. Nó thường gây ra sự bùng nổ bệnh cúm địa phương. Virut cúm C cũng chỉ lây nhiễm cho con người. Chúng lây nhiễm cho hầu hết con người khi họ có tuổi đời còn trẻ và hiếm khi gây ra các chứng bệnh trầm trọng.

Việc dùng vacxin tạo ra sự bảo vệ kháng lại bệnh do một tác nhân tương tự gây ra bằng cách làm cho đối tượng tăng sự phòng thủ trước khi lây nhiễm. Thông thường, việc này đã được thực hiện thông qua việc sử dụng dạng giảm độc lực sống hoặc dạng bất hoạt nguyên vẹn của các tác nhân lây nhiễm làm chất gây miễn dịch. Để tránh nguy hiểm của việc sử dụng virut nguyên vẹn (như virut bị giết hoặc giảm độc lực) làm vacxin, protein virut tái tổ hợp, ví dụ cấu trúc dưới đơn vị, đã được theo đuổi làm vacxin. Peptit và vacxin dưới đơn vị đều có nhiều giới hạn tiềm ẩn. Các vacxin dưới đơn vị có thể thể hiện tính gây miễn dịch kém, do việc gấp nếp không đúng đắn hoặc sự trình diện kháng nguyên kém. Vấn đề chính là khó để bảo đảm rằng sự thích ứng của protein được thiết kế ra là giống với sự thích ứng của kháng nguyên trong môi trường tự nhiên của chúng. Chất bổ trợ thích hợp và trong trường hợp peptit, protein vận chuyển, phải được sử dụng để làm tăng sự đáp ứng miễn dịch. Ngoài ra, các vacxin này gây ra các đáp ứng thể dịch lần đầu và do đó có

thể không gây ra được tính miễn dịch hữu hiệu. Vacxin dưới đơn vị thường là không hữu hiệu đối với các bệnh trong đó chứng minh được virut bất hoạt nguyên vẹn có thể tạo ra sự bảo vệ.

Hạt tương tự virut (virus-like particles-VLP) là ứng viên tiềm năng để bao gồm trong chế phẩm gây miễn dịch. VLP tương tự một cách chặt chẽ với virion trưởng thành, nhưng chúng không chứa vật liệu hệ gen virut. Do đó, VLP là không sao chép trong tự nhiên, mà làm cho chúng an toàn đối với việc dùng làm vacxin. Ngoài ra, VLP có thể được thiết kế để biểu hiện glycoprotein virut trên bề mặt của VLP, là cấu hình sinh lý nguyên thể nhất của chúng. Hơn nữa, do VLP tương tự với virion nguyên vẹn và là cấu trúc dạng hạt đa hoá trị, VLP có thể là hữu hiệu hơn trong việc gây ra kháng thể trung hoà với glycoprotein nhiều hơn so với kháng nguyên protein bao ngoài có thể hoà tan.

VLP đã được tạo ra ở thực vật (công bố đơn quốc tế số WO2009/009876; WO 2009/076778; WO 2010/003225; WO 2010/003235; WO 2011/03522; WO 2010/148511) và ở hệ côn trùng và động vật có vú (Noad, R. and Roy, P., 2003, *Trends Microbiol* 11: 438-44; Neumann et al., 2000, J. Virol., 74, 547-551). Latham and Galarza (2001, J. Virol., 75, 6154-6165) báo cáo sự hình thành của VLP cúm ở tế bào côn trùng bị lây nhiễm với baculovirut tái tổ hợp đồng biểu hiện hemagglutinin (HA), neuramindaza (NA), gen M1, và M2. Nghiên cứu này chứng minh rằng virion protein cúm tự lắp ráp khi đồng biểu hiện trong tế bào nhân chuẩn và protein nền M1 được đòi hỏi để tạo ra VLP. Gomez-Puertas et al., (1999, J. Gen. Virol, 80, 1635-1645) thể hiện rằng sự biểu hiện quá mức của M2 phong bế một cách hoàn toàn sự truyền dẫn ARN CAT vào môi trường nuôi cấy MDCK.

Glycoprotein hemagglutinin (HA) đỉnh của virut cúm là quan trọng nhất đối với sự hấp thu hạt virut bởi tế bào vật chủ. Nó chịu trách nhiệm đối với việc gắn của chúng vào các thụ thể tế bào chứa axit sialic và có liên quan đến việc thẩm thấu virut qua việc dung hợp vỏ bọc virut với màng sinh chất. Hoạt tính dung hợp và tiếp theo là khả năng lây nhiễm virut phụ thuộc vào việc tách phân tử tiền chất HA, HA0, thành chuỗi polypeptit liên kết disulfua, HA1 và HA2. Việc tách sự thay đổi hình dạng phụ thuộc độ pH tiếp theo

dẫn đến sự phơi trần và định vị lại peptit kỵ nước được bảo tồn ở mức cao ở đầu cuối amino của polypeptit HA2 chuyển qua màng, mà qua trung gian dung hợp màng.

HA được tổng hợp như tiền chất protein HA0, mà trải qua việc xử lý phân giải protein thành hai cấu trúc dưới đơn vị (HA1 và HA2) liên kết cùng với nhau bằng cầu disulfua. Hai đặc điểm cấu trúc được cho là có liên quan đến khả năng tách HA: trong các HA có khả năng tách bị giới hạn, đoạn liên kết thường gồm có arginin đơn lẻ, trong khi khả năng tách của HA nằm trong khoảng rộng các loại tế bào khác nhau có sự xen vào của một dãy gốc đa bazơ ở vị trí này với motif nhận biết enzym chính Arg-X-Lys/Arg-Arg, trong đó X là axit amin không bazơ. HA có vị trí tách nhiều bazơ được tách dựa vào đường vận chuyển xuất bào trước khi chúng đến vị trí sinh chồi trên bề mặt tế bào, ngược với HA với cầu liên kết một bazơ mà được bất hoạt trên hạt virut trong khoảng trống ngoại bào hoặc, như được thể hiện đối với chủng WSN, ở giai đoạn đi vào của virut. Thể quyết định thứ hai của việc tách HA xuất hiện sẽ là chuỗi phụ carbohydrat mà có mặt trong vùng lân cận của vị trí tách và cản trở khả năng tiếp cận proteaza. Việc mất carbohydrat dẫn đến khả năng tách HA và khả năng sinh bệnh của virut gia tăng.

Các chủng virut cúm động vật có vú và không gây bệnh chim gây ra các lây nhiễm khu trú về mặt giải phẫu là kết quả của khoảng giới hạn tế bào tiết proteaza mà có thể tách tiền chất HA0 về mặt ngoại bào (Chen J, et. al. 1998, Cell. Vol 95 :409-417). Các proteaza đáp ứng đối với việc tách HA0 trong sự lây nhiễm cúm của con người, được tiết bởi tế bào của đường hô hấp hoặc bằng cách đồng lây nhiễm vi khuẩn hoặc mycoplasma hoặc chúng có thể được tạo ra trong các đáp ứng viên với sự lây nhiễm. Ứng viên proteaza chính là tryptaza Clara, mà được tạo ra bởi tế bào Clara của biểu mô phế quản và có sự phân bố mô bị giới hạn (đường hô hấp trên). Proteaza đặc hiệu đối với trình tự monobazơ Q/E-X-R được tìm thấy ở vị trí tách của H1, H2, H3, và H6. HA từ H9 và chủng B thể hiện vị trí tách monobzzơ khác nhau một cách không đáng kể lần lượt với trình tự SSR và KER. Không có proteaza nào được nhận dạng đối với phần lớn virut cúm mà gây ra việc sàng lọc lây nhiễm đường ruột và đường hô hấp ở chim nước. Hầu hết các dòng tế bào không hỗ trợ sự sao chép nhiều chu trình trừ phi proteaza ngoại sinh (thường là trypsin) được bổ sung vào.

Tuy nhiên, ở các chủng chim sinh bệnh ở mức cao, HA0 được tách bởi họ proteaza nội bào phổ biến hơn, dẫn đến sự lây nhiễm cúm hệ thống. Sự khác nhau về sinh bệnh này tương quan với sự khác nhau cấu trúc ở vị trí tách HA0. Các chủng sinh bệnh có sự xen vào của axit amin polybazơ nằm trong hoặc gần với, vị trí monobazơ. Trong trường hợp này, việc tách xảy ra về mặt nội bào và proteaza có liên quan đã được nhận dạng là furin, và các enzym tương tự subtilisin khác, được tìm thấy trong tế bào golgi và có liên quan đến việc xử lý sau dịch mã của hormon và tiền chất nhân tố sinh trưởng. Trình tự nhận biết furin R-X-R/K-R là axit amin xen vào thường xuyên ở vị trí tách HA0 của H5 và H7. Sự phân bố mô rộng của enzym và hiệu quả của việc tách nội bào, góp phần vào sự lây nhiễm hệ thống lan toả rộng và gây hại do các virut này gây ra.

Vị trí tách HA là đích để giảm độc lực virut, do việc tách hoạt hoá của tiền chất HA0 thành các mảnh HA1 và HA2 bởi proteaza vật chủ là bước trong chu trình sao chép của tất cả các chủng virut cúm A và B. Chỉ có HA được tách có thể trải qua sự thay đổi hình dạng trong môi trường axit của nội thể sau khi nhập bào qua trung gian thụ thể để phơi trần đầu cuối N kỵ nước của mảnh HA2 để làm trung gian sự dung hợp giữa các màng nội thể và virion.

Horimoto T, et.al. (2006, Vaccine, Vol 24 : 3669-3676) mô tả sự huỷ bỏ vị trí tách polybazơ của H5 (RERRRKKR↓G) trong H5. Các thể đột biến được chọn được cho vào nghiên cứu gây miễn dịch ở chuột, bao gồm thể đột biến với việc xóa bỏ 4 axit amin mang điện tích thứ nhất (RERR) và sự cải biến để làm bất hoạt vị trí tách polybazơ (RKKR với TETR). Việc huỷ bỏ vị trí tách không tác động đến các đặc tính gây miễn dịch của thể đột biến H5. Việc huỷ bỏ vị trí polybazơ (GERRRKKR↓G được thay thế bằng RETR) để tạo ra thể đột biến NIBSC 05/240 virut tham khảo cúm NIBSC NIBG-23, cũng đã được báo cáo. Hoffman et. al. (2002, 2002, Vaccine, Vol 20 :3165-3170) thay thế vị trí tách polybazơ của H5 HA bằng vị trí monobazơ của H6 để làm tăng sự biểu hiện ở trứng. Bốn gốc thứ nhất được loại bớt và thay thế bốn axit amin sau cùng của vị trí polybazơ bằng IETR (sự thay thế của RERRRKKR↓G bằng IETR↓G). Thể đột biến H5 này thể hiện mức biểu hiện ở mức cao, phân giải protein tiềm ẩn và sự thay đổi hình dạng ở độ pH thấp được đòi hỏi đối với sự sao chép virut và tạo ra ở tế bào vật chủ, các dữ liệu gây miễn dịch

là không được báo cáo. Các nghiên cứu này thể hiện rằng sự cải biến của vị trí tách có thể được dùng để làm giảm tính độc của hạt virut (trong trường hợp mà trong đó virut thật được sao chép), cho phép virut sao chép mà không cần phải giết trứng vật chủ. Không cần phải có đột biến này, virut giết trứng trước khi đạt đến độ chuẩn ở mức cao.

Công bố đơn quốc tế số WO2013043067 của Sirko et al. mô tả vacxin ADN dùng cho gà mà chứa ADN bổ trợ mã hoá H5 haemagglutinin (HA) protein được cải biến trong đó vị trí tách phân giải protein giữa các dưới đơn vị HA được loại bớt. Sirko et al. nêu ra rằng điều này tạo ra chiến lược lớn hơn của vacxin và sự biểu hiện của "siêu kháng nguyên" ở dạng polypeptit dài, không được xử lý. Sirko et al. còn nêu ra rằng vùng mã hoá của HA được cải biến theo cách mà sự tạo ra protein ở tế bào chim đạt được hiệu suất tối đa. Sự cải biến chính là sự tối ưu hoá codon đối với gà và loại bớt vùng nhạy phân giải protein của HA.

Công bố đơn quốc tế số WO 2013/044390 mô tả phương pháp tạo ra hạt tương tự virut (VLP) ở thực vật với hemagglutinin (HA) được cải biến trong đó protein HA được cải biến bao gồm vòng phân giải protein được cải biến. HA được cải biến được biểu hiện với sự có mặt của vùng điều hoà virut khảm cây đậu đũa (CPMV) HT và yếu tố khuếch đại geminivirut từ virut vàng lùn đậu (Bean Yellow Dwarf Virus-BeYDV).

Công bố đơn yêu cầu cấp patent Mỹ số US 2008/0069821 của Yang et al. mô tả polypeptit và các biến thể polynucleotit của cúm HA để sử dụng trong việc tạo ra virut cúm làm vacxin. Virut cúm tái sắp xếp thu được bằng cách đưa nhóm phụ vật truyền tương ứng vào đoạn hệ gen của virut cúm chủ, trong tổ hợp với các đoạn bổ trợ thu được từ các biến thể của cúm. Thông thường, các chủng chủ được chọn dựa vào các đặc tính mong muốn liên quan đến việc dùng vacxin. Ví dụ, để tạo ra vacxin, ví dụ để tạo ra vacxin sống giảm độc lực, chủng virut cho chủ có thể được chọn đối với kiểu hình giảm độc lực, sự thích ứng lạnh và/hoặc nhạy nhiệt độ.

Bản chất kỹ thuật của sáng chế

Sáng chế đề cập đến việc tạo ra hạt tương tự virut và protein HA được cải biến ở thực vật.

Mục đích của sáng chế là đề xuất việc tạo ra hạt tương tự virut và protein HA được cải thiện ở thực vật.

Được mô tả trong bản mô tả này là axit nucleic bao gồm vùng tăng cường sự biểu hiện có hoạt tính ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá hemagglutinin (HA) cúm được cải biến bao gồm vòng phân giải protein được cải biến.

Hơn nữa, được mô tả trong bản mô tả này là phương pháp (A) tạo ra hạt tương tự virut (VLP) ở thực vật bao gồm các bước:

a) đưa axit nucleic bao gồm vùng tăng cường sự biểu hiện có hoạt tính ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá hemagglutinin (HA) cúm được cải biến bao gồm vòng phân giải protein được cải biến vào trong thực vật, hoặc phần của thực vật;

b) ủ thực vật hoặc phần của thực vật trong các điều kiện mà cho phép sự biểu hiện của axit nucleic, do đó tạo ra VLP.

Ngoài ra, được mô tả trong bản mô tả này là phương pháp (B) để tạo ra hạt tương tự virut (VLP) cúm ở thực vật bao gồm các bước:

a) cung cấp thực vật hoặc phần của thực vật, bao gồm axit nucleic bao gồm vùng tăng cường sự biểu hiện có hoạt tính ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá hemagglutinin (HA) cúm được cải biến bao gồm vòng phân giải protein được cải biến, và

b) ủ thực vật hoặc phần của thực vật trong các điều kiện mà cho phép sự biểu hiện của axit nucleic, do đó tạo ra VLP.

Hơn nữa, được mô tả trong bản mô tả này phương pháp (C) tạo ra protein HA được cải biến bao gồm vòng phân giải protein được cải biến bao gồm một hoặc nhiều vị trí phân cắt bởi proteaza thể hiện việc tách giảm hoặc huỷ bỏ ở thực vật bao gồm các bước:

a) đưa axit nucleic bao gồm vùng tăng cường sự biểu hiện có hoạt tính ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá hemagglutinin (HA) cúm được cải biến bao gồm vòng phân giải protein được cải biến vào trong thực vật, hoặc phần của thực vật;

b) ủ thực vật hoặc phần của thực vật trong các điều kiện mà cho phép sự biểu hiện của protein HA, do đó tạo ra protein HA được cải biến,

c) thu hoạch thực vật và tinh chế protein HA được cải biến.

Các phương pháp (A), (B) hoặc (C) như được mô tả trên đây còn có thể bao gồm các bước:

c) thu hoạch thực vật, và

d) tinh chế VLP, trong đó VLP có kích cỡ nằm trong khoảng từ 80 đến 300nm.

Hơn nữa, được mô tả trong bản mô tả này là phương pháp (D) làm gia tăng sản lượng sản phẩm của protein HA ở thực vật, bao gồm các bước:

a) đưa axit nucleic bao gồm vùng tăng cường sự biểu hiện có hoạt tính ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá hemagglutinin (HA) cúm được cải biến bao gồm vòng phân giải protein được cải biến vào trong thực vật, hoặc phần của thực vật;

b) ủ thực vật hoặc phần của thực vật trong các điều kiện mà cho phép sự biểu hiện của protein HA, do đó tạo ra protein HA được cải biến,

c) thu hoạch thực vật và tinh chế protein HA.

Vùng tăng cường sự biểu hiện có thể là CPMVX, CPMVX+ hoặc CPMV-HT+. Hơn nữa, axit nucleotit có thể không bao gồm yếu tố khuếch đại geminivirut. Do đó, axit nucleotit có thể không bao gồm vùng liên gen dài của virut vàng lùn đậu (BeYDV LIR), và vùng liên gen ngắn BeYDV (BeYDV SIR). Vòng phân giải protein được cải biến có thể bao gồm một hoặc nhiều vị trí phân cắt bởi proteaza biểu hiện sự phân cắt bởi proteaza

bị giảm hoặc bị hủy bỏ. Proteaza có thể là tương tự Clara hoặc tương tự Furin. Hơn nữa, vòng phân giải protein được cải biến có thể bao gồm trình tự liên kết và trình tự liên kết này có thể có trình tự axit amin GG (ví dụ, các vị trí axit amin 341-342 của SEQ ID NO: 17), TETQ (ví dj, các vị trí axit amin 9-12 của SEQ ID NO: 46) hoặc TETR (ví dụ, các vị trí axit amin 9-12 của SEQ ID NO: 45). HA được cải biến có thể bao gồm peptit tín hiệu nguyên thể hoặc không nguyên thể. Hơn nữa, trình tự nucleotit mã hoá HA được cải biến có thể bao gồm trình tự nucleotit khảm mã hoá, theo dãy, HA ectodomain được cải biến bao gồm vòng phân giải protein được cải biến, miền xuyên màng cúm, và đuôi tế bào chất, trong đó HA ectodomain được cải biến là từ chủng cúm thứ nhất và miền xuyên màng và đuôi tế bào chất là từ chủng cúm thứ hai.

Vòng phân giải protein được cải biến có thể bao gồm một hoặc nhiều vị trí phân cắt bởi proteaza biểu hiện sự phân cắt bởi proteaza bị giảm hoặc bị hủy bỏ. Hơn nữa, trình tự nucleotit mã hoá HA được chọn từ nhóm gồm có B HA, C, H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, và H16. Cũng được mô tả trong bản mô tả này là hạt tương tự virut (VLP) được tạo ra bằng phương pháp (A). Hạt tương tự virut (VLP) có thể bao gồm N-glycan đặc hiệu thực vật hoặc N-glycan được cải biến.

Sáng chế cũng đề xuất chế phẩm bao gồm liều hữu hiệu của VLP để gây ra đáp ứng miễn dịch, và chất mang dược dụng.

Cũng được mô tả trong bản mô tả này là axit nucleic bao gồm trình tự nucleotit mã hoá hemagglutinin (HA) cúm, trình tự nucleotit được liên kết theo cách hoạt động được với vùng điều hoà mà có hoạt tính ở thực vật, trong đó HA bao gồm trình tự vòng phân giải protein được cải biến. Axit nucleic có thể mã hoá HA bao gồm vòng phân giải protein được cải biến, trong đó protein có hoạt tính hemagglutinin (HA). Thực vật bao gồm axit nucleic cũng được đề xuất. Cũng được bao gồm là hạt tương tự virut (VLP) được tạo ra ở thực vật, VLP bao gồm hemagglutinin (HA) virut cúm được mã hoá bởi axit nucleic và một hoặc nhiều hơn một lipit thu được từ thực vật.

Phần bản chất kỹ thuật của sáng chế không nhất thiết phải mô tả tất cả các dấu hiệu của sáng chế.

Mô tả vắn tắt các hình vẽ

Các dấu hiệu này và các dấu hiệu khác của sáng chế sẽ trở nên rõ ràng hơn dựa vào phần mô tả sau đây, trong đó việc tham khảo được thực hiện đến các hình vẽ kèm theo trong đó:

Fig.1 thể hiện các thành phần được sử dụng để tạo ra A-2X35S/CPMV-HT/H5 Indonesia/NOS (Cấu trúc số 489). Fig.1A thể hiện đoạn mồi IF-H5A-I-05.s1+3c (SEQ ID NO: 2). Fig.1B thể hiện đoạn mồi IF-H5dTm.r (SEQ ID NO: 3). Fig.1C thể hiện sự biểu diễn theo sơ đồ của cấu trúc 1191. Fig.1D thể hiện cấu trúc 1191 (SEQ ID NO 4). Fig.1E thể hiện cátxét biểu hiện số 489 (SEQ ID NO 5). Fig.1F thể hiện trình tự axit amin của H5 từ cúm A/Indonesia/5/2005 (H5N1) (SEQ ID NO: 6). Fig.1G thể hiện trình tự nucleotit mã hoá H5 từ cúm A/Indonesia/5/2005 (H5N1) (SEQ ID NO: 42).

Fig.2 thể hiện các thành phần được sử dụng để tạo ra B-2X35S/CPMV HT/M2 New Caledonia/NOS (Cấu trúc số 1261). Fig.2A thể hiện đoạn mồi IF-S1-M1+M2ANC.c (SEQ ID NO:7). Fig.2B thể hiện đoạn mồi IF-S1-4-M2ANC.r (SEQ ID NO: 8). Fig.2C thể hiện trình tự nucleotit đối với gen M2 được tổng hợp (tương ứng với nt 1-26 được kết nối với 715-982 từ Số truy cập Genbank DQ508860) (SEQ ID NO: 9). Fig.2D thể hiện cátxét biểu hiện số 1261 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. M2 từ cúm A/New Caledonia/20/1999 (H1N1) được gạch dưới. (SEQ ID NO: 10). Fig.2E thể hiện trình tự axit amin của M2 từ cúm A/New Caledonia/20/1999 (H1N1) (SEQ ID NO: 11).

Fig.3 thể hiện thành phần được sử dụng để tạo ra C-2X35S/CPMV-HT/M2 Puerto Rico/NOS (Cấu trúc số 859). Fig.3A thể hiện trình tự nucleotit của gen M2 được tổng hợp (tương ứng với nt 26-51 được kết nối với nt 740-1007 từ số truy cập Genebank EF467824) (SEQ ID NO: 12). Fig.3B thể hiện cátxét biểu hiện số 859 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. M2 từ cúm A/Puerto Rico/8/1934 (H1N1) được gạch dưới. (SEQ ID NO: 13). Fig.3C thể hiện trình tự axit amin của M2 từ cúm A/Puerto Rico/8/1934 (H1N1) (SEQ ID NO:14).

Fig.4 thể hiện thành phần được sử dụng để tạo ra G-2X35S/CPMV-HT/PDISP/HA B Brisbane/NOS vào trong hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1008). Fig.4A

thể hiện sự biểu diễn theo sơ đồ của cấu trúc 1194. Các vị trí enzym giới hạn SacII và StuI được sử dụng để tuyến tính hoá plasmit được chú giải đối với sự biểu diễn. Fig.4B thể hiện cấu trúc 1194 từ phía bên trái đến bên phải các đường biên t-ADN (không gạch dưới). 2X35S/CPMV-HT/PDISP/NOS vào trong hệ khuếch đại BeYDV+Replicaza với Plastocyanin-P19-Plastoxyanin làm câm cátxét biểu hiện chất ức chế (SEQ ID NO: 31). Fig.4C thể hiện cátxét biểu hiện số 1008 từ BeYDV LIR bên trái sang BeYDV LIR bên phải. PDISP/HA từ cúm B/Brisbane/60/2008 được gạch dưới. (SEQ ID NO: 32).

Fig.5 thể hiện thành phần được sử dụng để tạo ra I-2X35S/CPMV-HT/PDISP/HA B Brisbane với vòng phân giải protein được xóa bỏ/NOS vào trong hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1059). Fig.5A thể hiện đoạn mồi 1039+1059.r (SEQ ID NO: 38). Fig.5B thể hiện đoạn mồi 1039+1059.c (SEQ ID NO: 39). Fig.5C thể hiện cátxét biểu hiện số 1059 từ BeYDV LIR bên trái sang BeYDV LIR bên phải. PDISP/HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ được gạch dưới (SEQ ID NO: 40). Fig.5D thể hiện trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ (SEQ ID NO: 41). Fig.5E thể hiện trình tự nucleotit của PDISP/HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ (SEQ ID NO: 43).

Fig.6 thể hiện thành phần được sử dụng để tạo ra B-2X35S/CPMV-HT/HA B Wisconsin/NOS vào trong hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1462). Fig.6A thể hiện đoạn mồi IF-HAB110.S1+3c (SEQ ID NO: 49). Fig.6B thể hiện đoạn mồi IF-HAB110.s1-4r (SEQ ID NO: 50). Fig.6C thể hiện trình tự nucleotit của HA B Wisconsin được tổng hợp (Số truy cập Genbank JN993010) (SEQ ID NO: 51). Fig.6D thể hiện sự biểu diễn theo sơ đồ của cấu trúc 193. Fig.6E thể hiện cấu trúc 193 từ bên trái sang bên phải đường biên t-ADN (gạch dưới). 2X35S/CPMV-HT/NOS vào trong hệ khuếch đại BeYDV(m)+Replicaza với Plastocyanine-P19-Plastocyanine cátxét biểu hiện chất ức chế (SEQ ID NO: 52). Fig.6F thể hiện trình tự nucleotit của cátxét biểu hiện số 1462 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. HA từ cúm B/Wisconsin/1/2010 được gạch chân (SEQ ID NO: 53). Fig.6G thể hiện trình tự axit amin của HA từ cúm B/Wisconsin/1/2010 (SEQ ID NO: 54). Fig.6H thể hiện sự biểu diễn theo sơ đồ của cấu trúc 1462.

Fig.7 thể hiện thành phần được sử dụng để tạo ra C-2X35S/CPMV-HT/HA B Wisconsin với vòng phân giải protein được xóa bỏ/NOS vào trong hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1467). Fig.7A thể hiện đoạn mồi HAB110(PrL-).r (SEQ ID NO: 55). Fig.7B thể hiện đoạn mồi HAB110(PrL-).c (SEQ ID NO: 56). Fig.7C thể hiện trình tự nucleotit của cátxét biểu hiện số 1467 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. HA từ cúm B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ được gạch chân (SEQ ID NO: 57). Fig.7D thể hiện trình tự axit amin của cúm B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ (SEQ ID NO: 58). Fig.7E thể hiện sự biểu diễn theo sơ đồ của cấu trúc 1467.

Fig.8 thể hiện thành phần được sử dụng để tạo ra A-2X35S/CPMV-HT/PDISP/HA B Brisbane với vòng phân giải protein được xóa bỏ/NOS (Cấu trúc số 1039). Fig.8A thể hiện trình tự nucleotit của cátxét biểu hiện số 1039 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ được gạch dưới (SEQ ID NO: 15). Fig.8B thể hiện sự biểu diễn theo sơ đồ của cấu trúc 1039.

Fig.9 thể hiện bản đồ plasmit của cấu trúc số 1008. Cấu trúc số 1008 dẫn sự biểu hiện của HA kiểu hoang dại từ chủng cúm B/Brisbane/60/2008. Cấu trúc này bao gồm các yếu tố thu được từ BeYDV để khuếch đại ADN.

Fig.10 thể hiện bản đồ plasmit của cấu trúc số 1059. Cấu trúc số 1059 dẫn sự biểu hiện của thể đột biến HA từ chủng cúm B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ. Cấu trúc này bao gồm yếu tố thu được từ BeYDV để khuếch đại ADN.

Fig.11 thể hiện bản đồ plasmit của cấu trúc số 1261. Cấu trúc số 1261 dẫn sự biểu hiện của M2 kiểu hoang dại từ chủng cúm A/New Caledonia/20/99 (H1N1).

Fig.12 thể hiện bản đồ plasmit của cấu trúc số 859. Cấu trúc số 859 dẫn sự biểu hiện của M2 kiểu hoang dại từ chủng cúm A/Puerto Rico/8/34 (H1N1).

Fig.13A thể hiện phép phân tích thấm tách Western của sự biểu hiện protein HA ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. "1008": sự biểu hiện của HA kiểu hoang dại từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV); "1008+1261":

sự đồng biểu hiện của HA kiểu hoang dại từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2 từ A/New Caledonia/20/99; "1059": sự biểu hiện của thể đột biến HA từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV); "1059+1261": sự đồng biểu hiện của thể đột biến HA từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2 từ A/New Caledonia/20/99. Các thực vật từ ba lần thấm lọc riêng biệt được phân tích (A, B và C). Tỷ lệ cho biết rằng tỷ phần của môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm đồn biểu hiện. Fig.13B thể hiện phép so sánh về khả năng hemagglutin hoá của phần chiết protein thô từ thực vật tạo ra HA.

Fig.14 thể hiện phép phân tích thấm tách Western của sự biểu hiện protein HA ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. "1059": sự biểu hiện của thể đột biến HA từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV); "1059+1261": sự đồng biểu hiện của thể đột biến HA từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2 từ A/New Caledonia/20/99. "1059+859": sự đồng biểu hiện của thể đột biến HA từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2 từ A/Puerto Rico/8/34. Các thực vật từ ba lần thấm lọc riêng biệt được phân tích (A, B và C). Các tỷ lệ cho biết tỷ phần của môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm đồng nuôi cấy.

Fig.15 thể hiện sự căn thẳng hàng trình tự axit amin của vùng bao quanh cầu liên kết của HA từ vài chủng cúm: H1 New Cal (SEQ ID NO:22); H1 Brisbane (SEQ ID NO:23); H1 Sol Islands (SEQ ID NO:24); H2A Singapore (SEQ ID NO:25); H3A Brisbane (SEQ ID NO:26); H3A WCN (SEQ ID NO:27); H5 Anhui (SEQ ID NO:28); H5 Indo (SEQ ID NO:29); H5 Vietnam (SEQ ID NO:30); H6 Teal HK (SEQ ID NO:33); H7 Eq Prague (SEQ ID NO:34); H9A HK (SEQ ID NO:35); B Florida (SEQ ID NO:36); B Malaysia (SEQ ID NO:37). Vị trí tách của tiền chất HA0 được chỉ ra bằng mũi tên.

Fig.16A thể hiện phép phân tích thấm tách Western của sự biểu hiện protein HA ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. HA từ B/Wisconsin/1/2010 được đồng biểu hiện với M2 từ A/New Caledonia/20/99. Mười microgam phần chiết protein được nạp vào mỗi dãy. "C+": đối chứng dương tính, B/Wisconsin/1/2010 virut bán tinh chế từ

National Institute for Biological Standards and Control, United Kingdom; "1462": sự biểu hiện của HA kiểu hoang dại từ B/Wisconsin/1/2010 với sự có mặt của yếu tố khuếch đại (BeYDV); "1467": sự biểu hiện của thể đột biến HA từ B/Wisconsin/1/2010 với sự có mặt của yếu tố khuếch đại (BeYDV); "1462+1261": sự đồng biểu hiện của HA kiểu hoang dại từ B/Wisconsin/1/2010 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2; "1467+1261": sự đồng biểu hiện của thể đột biến HA từ B/Wisconsin/1/2010 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2. Các tỷ lệ cho biết mật độ quang học đối với mỗi môi trường nuôi cấy *Agrobacterium* được sử dụng trong các thử nghiệm biểu hiện và đồng thử nghiệm. Fig.16B thể hiện phép so sánh khả năng hemagglutin hoá của phần chiết protein thô từ các thực vật được biến nạp với AGL1/1462, AGL1/1467, AGL1/1462+AGL1/1261 và AGL1/1467+AGL1/1261.

Fig.17A và Fig.17B thể hiện phép phân tích thấm tách Western của sự biểu hiện protein HA ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. "1008": sự biểu hiện của HA kiểu hoang dại từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV); "1008+1261": sự đồng biểu hiện của HA kiểu hoang dại từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2 từ A/New Caledonia/20/99;"1039": sự biểu hiện của thể đột biến HA từ B/Brisbane/60/2008 với sự không có mặt của yếu tố khuếch đại (BeYDV). "1039+1261": sự đồng biểu hiện của thể đột biến HA từ B/Brisbane/60/2008 với sự không có mặt của yếu tố khuếch đại (BeYDV) với M2 từ A/New Caledonia/20/99.HA từ A/Brisbane/59/2007 (H1N1). "1059": sự biểu hiện của thể đột biến HA từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV); "1059+1261": sự đồng biểu hiện của thể đột biến HA từ B/Brisbane/60/2008 với sự có mặt của yếu tố khuếch đại (BeYDV) với M2 từ A/New Caledonia/20/99.

Fig.18A thể hiện phép biểu diễn theo sơ đồ của việc tách HA0 bởi proteaza tương tự Clara và/hoặc tương tự Furin thành HA1 và HA2. Fig.18B thể hiện việc căn thẳng hàng trình tự của HA từ vài chủng cúm: H1 New Cal (SEQ ID NO:22); H1 Brisbane (SEQ ID NO:23); H1 Sol Islands (SEQ ID NO:24); H2A Singapore (SEQ ID NO:25); H3A Brisbane (SEQ ID NO:26); H3A WCN (SEQ ID NO:27); H5 Anhui (SEQ ID NO:28); H5 Indo (SEQ ID NO:29); H5 Vietnam (SEQ ID NO:30); H6 Teal HK (SEQ ID NO:33); H7

Eq Prague (SEQ ID NO:34); H9A HK (SEQ ID NO:35); B Florida (SEQ ID NO:36); B Malaysia (SEQ ID NO:37). Fig.18C thể hiện việc xóa bỏxóa bỏ một phần của vị trí tách ở chủng H5, A/Anhui/1/2005 (H5N1) SEQ ID NO: 69, A/Indonesia/5/2005 (H5N1) SEQ ID NO: 70, A/Vietnam/1194/2004 (H5N1) SEQ ID NO: 71, và chủng typ B, B/Florida/4/2006 SEQ ID NO: 72, và B/Malaysia/2506/2004 SEQ ID NO:73.

Fig.19 thể hiện sự đột biến của vị trí tách trong H5/Indo. Trình tự nguyên thể (SEQ ID NO: 44); H5/Indo cải biến vị trí tách bao gồm TETR (SEQ ID NO:45); H5/Indo cải biến vị trí tách bao gồm TETQ (SEQ ID NO:46).

Fig.20 thể hiện độ chuẩn được tìm thấy trong sinh khối ban đầu sau khi chiết enzym của các H5/Indo HA khác nhau bao gồm các đột biến nằm trong vòng phân giải protein. H5/Indo đối chứng (cấu trúc 489); vòng phân giải protein H5/Indo được thay thế bằng cầu liên kết GG (cấu trúc 928); vòng phân giải protein H5/Indo được thay thế bằng cầu liên kết TETR (cấu trúc 676); vòng phân giải protein H5/Indo được thay thế bằng cầu liên kết TETQ (cấu trúc 766).

Fig.21 thể hiện các phương pháp khác nhau để cải biến vòng phân giải protein chứa typ B HA. Fig.21A thể hiện trình tự axit amin của B/Brisbane/60/2008 (SEQ ID NO: 16) nguyên thể. Phần không gạch dưới là vòng phân giải protein và miền HA2. Fig.21B thể hiện trình tự axit amin của B/Brisbane/60/2008 với vòng phân giải protein cải biến (SEQ ID NO: 17), trong đó 19 gốc axit amin bao gồm trình tự AKLLKERGFFGAIAGFLEG được thay thế bằng cầu liên kết GG (chữ nghiêng). Fig.21C thể hiện trình tự axit amin của B/Brisbane/60/2008 (SEQ ID NO: 18), trong đó 9 axit amin bao gồm trình tự PPAKLLKER được thay thế bằng cầu liên kết –GSSSGSSSG- (chữ nghiêng). Fig.21D thể hiện trình tự axit amin của H3 A/Perth/16/2009 (SEQ ID NO: 19) nguyên thể. Fig.21E thể hiện trình tự axit amin của H3 A/Perth/16/2009 (SEQ ID NO: 20) với 12 gốc axit amin bao gồm trình tự RNVPEKQTRGIF được thay thế bằng cầu liên kết GS (chữ nghiêng). Fig.21F thể hiện trình tự axit amin của H3 A/Perth/16/2009 (SEQ ID NO: 21) với 9 gốc axit amin bao gồm trình tự RNVPEKQTR được thay thế bằng cầu liên kết GSSGSSGSS- (chữ nghiêng).

Fig.22 thể hiện phép phân tích thấm tách Western của sự biểu hiện HA ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. HA từ H5/Indo. Panen phía trên liệt kê các thành phần được nạp vào mỗi dãy; panen phía dưới thấm tách Western. C: H5 Indonesia/5/05 S-STD-0002 tái tổ hợp; kháng thể sơ cấp: kháng HA A/Indonesia/05/2005 CBER # S-BIO-0003 1/50 000; thấm tách. Các dãy 1 và 2: H5/Indo đối chứng (cấu trúc 489); các dãy 3 và 4 vòng phân giải protein H5/Indo được thay thế bằng cầu liên kết GG (cấu trúc 928); dãy 5 và vòng phân giải protein 6 H5/Indo được thay thế bằng cầu liên kết TETR (cấu trúc 676); các dãy 7 và 8 vòng phân giải protein H5/Indo được thay thế bằng cầu liên kết TETQ (cấu trúc 766); dãy 9 chất đánh dấu MW; dãy 10 H5 đối chứng.

Fig.23 thể hiện thành phần được sử dụng để tạo ra B-2X35S/CPMV-HT/H5 từ A/Indonesia/5/2005 với đột biến vị trí tách TETR (Cấu trúc số 676). Fig.23A thể hiện trình tự đoạn mồi MutCleavage-H5(Indo).r (SEQ ID NO:74). Fig.23B thể hiện trình tự đoạn mồi MutCleavage-H5(Indo).c (SEQ ID NO:75). Fig.23C thể hiện trình tự nucleotit (SEQ ID NO: 76) đối với cátxét biểu hiện 676 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. H5 từ cúm A/Indonesia/5/2005 (H5N1) thể đột biến vị trí tách TETR được gạch dưới. Fig.23D thể hiện trình tự axit amin (SEQ ID NO:77) của thể đột biến vị trí tách TETR của H5 từ cúm A/Indonesia/5/2005 (H5N1). Fig.23E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 676.

Fig.24 thể hiện thành phần được sử dụng để tạo ra B-2X35S/CPMV-HT/H5 từ A/Indonesia/5/2005 với đột biến vị trí tách TETQ (Cấu trúc số 766). Fig.24A thể hiện trình tự đoạn mồi H5I505_TETQ.r (SEQ ID NO:78). Fig.24B thể hiện trình tự đoạn mồi H5I505_TETQ.c (SEQ ID NO:79). Fig.24C thể hiện trình tự nucleotit (SEQ ID NO: 80) đối với cátxét biểu hiện 766 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. H5 từ cúm A/Indonesia/5/2005 (H5N1) thể đột biến vị trí tách TETQ được gạch dưới. Fig.24D thể hiện trình tự axit amin (SEQ ID NO:81) của thể đột biến vị trí tách TETQ của H5 từ cúm A/Indonesia/5/2005 (H5N1). Fig.24E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 766.

Fig.25 thể hiện thành phần được sử dụng để tạo ra B-2X35S/CPMV-HT/H5 từ A/Indonesia/5/2005 với vòng phân giải protein được xóa bỏ (Cấu trúc số 928). Fig.25A

thể hiện trình tự đoạn mồi H5I505(PrL-).r (SEQ ID NO: 82). Fig.25B thể hiện trình tự đoạn mồi H5I505(PrL-).c (SEQ ID NO: 83). Fig.25C thể hiện trình tự nucleotit (SEQ ID NO:84) đối với cátxét biểu hiện 928 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. H5 từ cúm A/Indonesia/5/2005 (H5N1) vòng phân giải protein được xóa bỏ được gạch dưới. Fig.25D thể hiện trình tự axit amin (SEQ ID NO:85) của thể đột biến của H5 từ cúm A/Indonesia/5/2005 (H5N1) bao gồm vòng phân giải protein được xóa bỏ. Fig.25E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 928.

Fig.26A thể hiện sơ đồ chung về ví dụ của vài trình tự tăng cường, CPMVX, và CPMVX+ (bao gồm CPMVX, và đoạn nhồi, mà trong ví dụ không hạn chế này, bao gồm vị trí tạo đa dòng và trình tự kozak thực vật), như được mô tả trong bản mô tả này. Mỗi CPMCX và CPMVX+ được thể hiện như được liên kết theo cách hoạt động được với vùng điều hoà thực vật ở các đầu 5' của chúng và ở các đầu 3' của chúng, theo dãy, trình tự nucleotit quan tâm (bao gồm vị trí khơi mào ATG và vị trí STOP), 3'UTR, và trình tự kết thúc. Ví dụ về cấu trúc CPMVX như được mô tả trong bản mô tả này, là CPMV160. Ví dụ về cấu trúc CPMVX+ như được mô tả trong bản mô tả này, là CPMV160+. Fig.26B thể hiện độ chuẩn hemagglutin hoá tương đối (HMG) trong phần chiết protein thô của protein HA được cải biến được tạo ra ở thực vật bao gồm các cấu trúc biểu hiện CPMV-HT và cấu trúc biểu hiện dựa trên CPMV160+. Các dữ liệu đối với sự biểu hiện của HA B Brisbane/60/08 với vòng phân giải protein được xóa bỏ và với peptit tín hiệu PDI (cấu trúc số 1039, 5'UTR: CMPV HT; và cấu trúc số 1937, 5'UTR: CMPV160+; xem ví dụ 5.7), B Brisbane/60/08+H1Tm, với vòng phân giải protein được xóa bỏ, với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1, B Massachusetts/2/2012 2012 với vòng phân giải protein được xóa bỏ và với peptit tín hiệu PDI (cấu trúc số 2072, 5'UTR: CMPV HT; và cấu trúc số 2050, 5'UTR: CMPV160+; xem ví dụ 5.14), B Massachusetts/2/2012+H1Tm với vòng phân giải protein được xóa bỏ, với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng của H1 A/California/07/2009 và với peptit tín hiệu PDI (cấu trúc số 2074, 5'UTR: CMPV HT; và cấu trúc số 2060, 5'UTR: CMPV160+; xem ví dụ 5.15), B Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ và với peptit tín hiệu nguyên thể (cấu trúc số 1445, 5'UTR: CMPV HT; và cấu trúc số 1975, 5'UTR: CMPV160+; xem ví dụ 5.16), và B Wisconsin/1/2010+H1Tm với vòng

phân giải protein được xóa bỏ, với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng của H1 A/California/07/2009 và với peptit tín hiệu nguyên thể (cấu trúc số 1454, 5'UTR: CMPV HT; và cấu trúc số 1893, 5'UTR: CMPV160+; xem ví dụ 5.18), là được thể hiện.

Fig.27A thể hiện sơ đồ chung của trình tự tăng cường của CPMV HT và CPMV HT+ được dung hợp với trình tự nucleotit quan tâm. Không phải tất cả các yếu tố được thể hiện trên Fig. này có thể được đòi hỏi nằm trong trình tự tăng cường. Các yếu tố bổ sung có thể được bao gồm ở đầu cuối 3' của trình tự nucleotit quan tâm (không được thể hiện trên hình vẽ) bao gồm trình tự mã hoá vùng không dịch mã 3' comovirut (UTR), plastocyanin 3' UTR, hoặc tổ hợp của comovirut 3' UTR và plastoxyanin 3' UTR. Fig.27B thể hiện độ chuẩn hemagglutin hoá tương đối (HMG) trong phần chiết protein thô của protein được tạo ra ở thực vật bao gồm cấu trúc biểu hiện CPMV-HT và cấu trúc biểu hiện dựa vào CPMV HT+, được liên kết theo cách hoạt động được với trình tự nucleotit quan tâm. Các dữ liệu để biểu hiện HA B Brisbane/60/08 với vòng phân giải protein được xóa bỏ và với peptit tín hiệu PDI (cấu trúc số 1039: CPMV HT; xem ví dụ 5.7 và cấu trúc số 1829: CPMV HT+; xem ví dụ 5.12), B Brisbane/60/08 + H1TM với vòng phân giải protein được xóa bỏ, với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1 A/California/07/2009, và với peptit tín hiệu PDI (cấu trúc số 1067: CPMV HT; xem ví dụ 5.14 và cấu trúc số 1875: CPMV HT+; xem ví dụ 5.19), B Massachusetts/2/2012 với vòng phân giải protein được xóa bỏ, và với peptit tín hiệu PDI (cấu trúc số 2072: CMPV HT; xem ví dụ 5.15 và cấu trúc số 2052: CMPV HT+; xem ví dụ 5.20), B Massachusetts/2/2012+H1Tm với vòng phân giải protein được xóa bỏ, với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1 A/California/07/2009 và với peptit tín hiệu PDI (cấu trúc số 2074: CMPV HT; xem ví dụ 5.16 và cấu trúc số 2062: CMPV HT+; xem ví dụ 5.21), B Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ và với peptit tín hiệu nguyên thể (cấu trúc số 1445: CMPV HT; xem ví dụ 5.17 và cấu trúc số 1839: CMPV HT+; xem ví dụ 5.22), và B Wisconsin/1/2010+H1Tm với vòng phân giải protein được xóa bỏ, với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1 A/California/07/2009 và với peptit tín

hiệu nguyên thể (cấu trúc số 1454: CMPV HT; xem ví dụ 5.18 và cấu trúc số 1860: CMPV HT+; xem ví dụ 5.23), là được thể hiện.

Fig.28A thể hiện phép phân tích thấm tách Western của sự biểu hiện H3 Perth protein ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. Dãy 1: (2019+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ H3 Perth-16-09 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT+) với M2 từ A/New Caledonia/20/99; Dãy 2: (2139+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ H3 Perth-16-09 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV 160+) với M2 từ A/New Caledonia/20/99; Dãy 3 (2039+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ H3 Perth-16-09 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV HT+) với M2 từ A/New Caledonia/20/99; Dãy 4: (2159+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ H3 Perth-16-09 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV 160+) với M2 từ A/New Caledonia/20/99.

Fig.28B thể hiện phép phân tích thấm tách Western của sự biểu hiện protein B Malaysia ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. Dãy 2: (2013+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ B Malaysia 2506-04 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-160+) với M2 từ A/New Caledonia/20/99; Dãy 2: (2014+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ B Malaysia 2506-04 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV 160+) với M2 từ A/New Caledonia/20/99.

Fig.28C thể hiện phép phân tích thấm tách Western của sự biểu hiện protein H9 Hong Kong ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. Dãy 1: (1610+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ H9 Hong Kong -1037-99 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT) với M2 từ A/New Caledonia/20/99; Dãy 2: (1630+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ H9 Hong Kong -1037-99 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT+) và yếu tố khuếch đại BEYDV với M2 từ A/New Caledonia/20/99; Dãy 3: (2244+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ H9 Hong Kong -1037-99 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT+) với M2 từ A/New Caledonia/20/99; Dãy

4: (2226+1261): sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ H9 Hong Kong -1037-99 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV 160+) với M2 từ A/New Caledonia/20/99. Dãy 6: (2246+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ H9 Hong Kong-1037-99 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-160+) và yếu tố khuếch đại BeYDV với M2 từ A/New Caledonia/20/99; Dãy 7: (2225+1261) sự đồng biểu hiện của HA đột biến (được cải biến) from H9 Hong Kong -1037-99 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT+) với M2 từ A/New Caledonia/20/99; Dãy 8: (2245+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ H9 Hong Kong -1037-99 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV HT+) và yếu tố khuếch đại BeYDV với M2 từ A/New Caledonia/20/99. Dãy 9: (2227+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ H9 Hong Kong -1037-99 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV 160+) với M2 từ A/New Caledonia/20/99. Dãy 10: (2247+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ H9 Hong Kong -1037-99 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV 160+) và yếu tố khuếch đại BeYDV với M2 từ A/New Caledonia/20/99.

Fig.28D thể hiện phép phân tích thấm tách Western của sự biểu hiện protein B Massachusetts ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. Dãy 1: (2070+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ B Massachusetts -2-12 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT) với M2 từ A/New Caledonia/20/99; Dãy 2: (2080+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ B Massachusetts -2-12 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT+) với M2 từ A/New Caledonia/20/99; Dãy 3: (2090+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ B Massachusetts -2-12 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-160+) với M2 từ A/New Caledonia/20/99; Dãy 4: (2072+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ B Massachusetts -2-12 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV HT) với M2 từ A/New Caledonia/20/99; Dãy 5: (2052+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ B Massachusetts -2-12 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV HT+) với M2 từ A/New Caledonia/20/99; Dãy 6: (2050+1261) sự đồng biểu hiện của HA đột biến (được cải biến)

từ B Massachusetts -2-12 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV 160+) với M2 từ A/New Caledonia/20/99.

Fig.28E thể hiện phép phân tích thấm tách Western của sự biểu hiện H2 Sin ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. Dãy 1: (2220+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ H2 Singapore-1-57 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT+) với M2 từ A/New Caledonia/20/99; Dãy 2: (2222+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ H2 Singapore -1-57 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV 160+) với M2 từ A/New Caledonia/20/99. Dãy 3: (2221+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ H2 Singapore -1-57 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT+) với M2 từ A/New Caledonia/20/99; Dãy 4: (2223+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ H2 Singapore-1-57 với sự có mặt của vùng tăng cường sự biểu hiện (CPMV 160+) với M2 từ A/New Caledonia/20/99.

Fig.28F thể hiện phép phân tích thấm tách Western của sự biểu hiện protein B/Forida ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp. Dãy 1: (1004+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ B/Florida với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT) với M2 từ A/New Caledonia/20/99; Dãy 2: (1003+1261) sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) từ B/Florida với sự có mặt của vùng tăng cường sự biểu hiện (CPMV HT) và yếu tố khuếch đại BeYDV với M2 từ A/New Caledonia/20/99. Dãy 3: (2102+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ B/Florida với sự có mặt của vùng tăng cường sự biểu hiện (CPMV-HT+) với M2 từ A/New Caledonia/20/99; Dãy 4: (2104+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ B/Florida với sự có mặt của vùng tăng cường sự biểu hiện (CPMV HT+) và yếu tố khuếch đại BeYDV với M2 từ A/New Caledonia/20/99. Dãy 5: (2106+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ B/Florida+H1 California TMCT với sự có mặt của vùng tăng cường sự biểu hiện (CPMV HT+) với M2 từ A/New Caledonia/20/99.Lane 6: (2108+1261) sự đồng biểu hiện của HA đột biến (được cải biến) từ B/Florida+H1 California TMCT với sự có mặt của vùng tăng cường sự biểu hiện (CPMV HT+) và yếu tố khuếch đại BeYDV với M2 từ A/New Caledonia/20/99.

Fig.29A thể hiện độ chuẩn HA tương đối của HA được cải biến từ các chủng cúm khác nhau mà được biểu hiện với sự có mặt của yếu tố tăng cường CPMV HT, CPMV HT+ hoặc CPMV 160+. Hoạt tính được so với protein HA nguyên thể được biểu hiện với cùng yếu tố tăng cường. H3 A Perth/16/09 (H3 Per1609), H3 Victoria/361/11 (H3 Vic26111), B Brisbane 60/2008 (HB Bris60008), B Malaysia 2506/04 (HB Mal 250604) và B Massachusetts 2/12 (Ma212) được đồng biểu hiện với M2 protein cúm.

Fig. 30 thể hiện các thành phần được sử dụng để điều chế F-2X35S/CPMV-HT/PDISP/HA B Brisbane/NOS (số Construct 1029).Fig.30A thể hiện đoạn mồi IF-S2+S4-B Bris.c (SEQ ID NO: 86). Fig.30B thể hiện đoạn mồi IF-S1a4-B Bris.r (SEQ ID NO: 87). Fig.30C thể hiện trình tự nucleotit của gen HA B Brisbane được tổng hợp (tương ứng với nt 34-1791 từ Số truy cập Genbank FJ766840) (SEQ ID NO: 88). Fig.30D thể hiện trình tự nucleotit của cátxét biểu hiện số 1029 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. PDISP/HA từ cúm B/Brisbane/60/2008 được gạch dưới. (SEQ ID NO: 89). Fig.30E thể hiện trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008 (SEQ ID NO: 90). Fig.30F thể hiện sự biểu diễn theo sơ đồ của cấu trúc 1029. Các vị trí enzym giới hạn SacII và StuI được sử dụng để tuyến tính hoá plasmit được chú giải đối với phép biểu diễn này.

Fig.31 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1039 và 1829 (2X35S/CPMV HT PDISP/HA B Brisbane (PrL-) NOS và 2X35S/CPMV HT+ PDISP/HA B Brisbane (PrL-) NOS, một cách tương ứng; xem ví dụ 5.12). Cấu trúc số 1039 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (PDISP/HA B Brisbane (PrL-)). Cấu trúc số 1829 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi bao gồm M protein không hoàn thiện, vị trí tạo đa dòng và trình tự kozak của thực vật là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV HT+. PDISP: peptit tín hiệu protein disulfua isomeraza; NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ. Fig.31A thể hiện trình tự nucleotit của PDISP/HA B Brisbane (PrL-) (SEQ ID NO: 91). Fig.31B thể hiện trình tự axit amin của

PDISP/HA B Brisbane (PrL-); SEQ ID NO: 92). Fig.31C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1829 (2X35S/CPMV HT+).

Fig.32 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1039 và 1937 (2X35S/CPMV HT PDISP/HA B Brisbane (PrL-) NOS và 2X35S/CPMV160+ PDISP/HA B Brisbane (PrL-) NOS, một cách tương ứng; xem ví dụ 5.7). Cấu trúc số 1039 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (PDISP/HA B Brisbane (PrL-)). Cấu trúc số 1937 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi (vị trí tạo đa dòng), và trình tự kozak của thực vật (cấu trúc này không bao gồm trình tự mã hoá M protein không hoàn thiện) và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV160+ (CPMVX+, trong đó X=160). PDISP: peptit tín hiệu protein disulfua isomeraza; NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ. Fig.32A thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1937 (2X35S/CPMV160+; cấu trúc được tạo thành chủ yếu từ CPMVX+, trong đó X=160).

Fig.33 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1067 và 1977 (2X35S/CPMV HT PDISP/HA B Brisbane (Prl-)+H1 California TMCT NOS và 2X35S/CPMV160+ PDISP/HA B Brisbane (PrL-)+H1 California TMCT NOS, một cách tương ứng; xem ví dụ 5.14). Cấu trúc số 1067 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (PDISP/HA B Brisbane (PrL-)+H1 California TMCT). Cấu trúc số 1977 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi (vị trí tạo đa dòng), và trình tự kozak của thực vật (cấu trúc này không bao gồm trình tự mã hoá M protein không hoàn thiện) và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV160+ (CPMVX+, trong đó X=160). PDISP: peptit tín hiệu protein disulfua isomeraza; NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ; TMCT: đuôi tế bào chất của miền xuyên màng. Fig.33A thể hiện trình tự nucleotit của PDISP/HA B Brisbane (PrL-)+H1 California TMCT (SEQ ID NO: 95). Fig.33B thể hiện

trình tự axit amin của PDISP/HA B Brisbane (PrL-)+H1 California TMCT (SEQ ID NO: 96). Fig.33C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1067 (2X35S/CPMV HT; cấu trúc tham khảo). Fig.33D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1977 (2X35S/CPMV160+; cấu trúc được tạo thành chủ yếu từ CPMVX+, trong đó X=160).

Fig.34 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2072 và 2050 (2X35S/CPMV HT PDISP/HA B Massachusetts (PrL-) NOS và 2X35S/CPMV160+ PDISP/HA B Massachusetts (PrL-) NOS, một cách tương ứng; xem ví dụ 5.15). Cấu trúc số 2072 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (PDISP/HA B Massachusetts (PrL-)). Cấu trúc số 2050 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi (vị trí tạo đa dòng), và trình tự kozak của thực vật (cấu trúc này không bao gồm trình tự mã hoá M protein không hoàn thiện) và là một ví dụ về cấu trúc được tạo thành chủ yếu từ cấu trúc được tạo thành chủ yếu từ CPMV160+ (CPMVX+, trong đó X=160). PDISP: peptit tín hiệu protein disulfua isomeraza; NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ. Fig.34A thể hiện trình tự nucleotit của PDISP/HA B Massachusetts (PrL-) (SEQ ID NO: 97). Fig.34B thể hiện trình tự axit amin của PDISP/HA B Massachusetts (PrL-) (SEQ ID NO: 98). Fig.34C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2072 (2X35S/CPMV HT; cấu trúc tham khảo). Fig.34D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2050 (2X35S/CPMV160+; cấu trúc được tạo thành chủ yếu từ CPMVX+, trong đó X=160).

Fig.35 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2074 và 2060 (2X35S/CPMV HT PDISP/HA B Massachusetts (PrL-)+H1 California TMCT NOS và 2X35S/CPMV160+ PDISP/HA B Massachusetts (PrL-)+H1 California TMCT NOS, một cách tương ứng; xem ví dụ 5.16). Cấu trúc số 2074 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (PDISP/HA B Massachusetts (PrL-)+H1 California TMCT). Cấu trúc số 2060 bao gồm CPMV 5'UTR bao gồm 160 nucleotit,

mảnh nhồi (vị trí tạo đa dòng), và trình tự kozak của thực vật (cấu trúc này không bao gồm trình tự mã hoá M protein không hoàn thiện) và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV160+ (CPMVX+, trong đó X=160). PDISP: peptit tín hiệu protein disulfua isomeraza; NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ; TMCT: đuôi tế bào chất của miền xuyên màng. Fig.35A thể hiện trình tự nucleotit của PDISP/HA B Massachusetts (PrL-)+H1 California TMCT (SEQ ID NO: 99). Fig.35B thể hiện trình tự axit amin của PDISP/HA B Massachusetts (PrL-)+H1 California TMCT (SEQ ID NO: 100). Fig.35C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2074 (2X35S/CPMV HT; cấu trúc tham khảo). Fig.35D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2060 (2X35S/CPMV160+; cấu trúc được tạo thành chủ yếu từ CPMVX+, trong đó X=160).

Fig.36 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1445, 1820 và 1975 (2X35S/CPMV HT HA B Wisconsin (PrL-) NOS, 2X35S/CPMV160+ HA B Wisconsin (PrL-) NOS và 2X35S/CPMV160 HA B Wisconsin (PrL-) NOS, một cách tương ứng; xem ví dụ 15.17). Cấu trúc số 1445 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (HA B Wisconsin (PrL-)). Cấu trúc số 1820 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi (vị trí tạo đa dòng), và trình tự kozak của thực vật (cấu trúc này không bao gồm trình tự mã hoá M protein không hoàn thiện) và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV160+ (CPMVX+, trong đó X=160). Cấu trúc số 1975 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, và không bao gồm mảnh nhồi (vị trí tạo đa dòng), hoặc trình tự kozak của thực vật (cấu trúc này cũng không bao gồm trình tự mã hoá M protein không hoàn thiện) và là một ví dụ về cấu trúc được tạo thành chủ yếu từ "CPMV160" (CPMVX). PrL-: vòng phân giải protein được xóa bỏ; NOS: đoạn kết thúc nopalin synthaza. Fig.36A thể hiện trình tự nucleotit của HA B Wisconsin (PrL-) (SEQ ID NO: 101). Fig.36B thể hiện trình tự axit amin của HA B Wisconsin (PrL-) (SEQ ID NO: 102). Fig.36C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1445 (2X35S/CPMV HT; cấu trúc tham khảo). Fig.36D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1820 (2X35S/CPMV160+; cấu trúc được tạo thành chủ yếu từ

CPMVX+). Fig.36E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1975 (2X35S/CPMV160; cấu trúc được tạo thành chủ yếu từ CPMVX, trong đó X=160).

Fig.37 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1454 và 1893 (2X35S/CPMV HT HA B Wisconsin (PrL-)+H1 California TMCT NOS và 2X35S/CPMV160+ HA B Wisconsin (PrL-)+H1 California TMCT NOS, một cách tương ứng; xem ví dụ 5.18). Cấu trúc số 1454 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (HA B Wisconsin (PrL-)+H1 California TMCT). Cấu trúc số 1893 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi (vị trí tạo đa dòng), và trình tự kozak của thực vật (cấu trúc này không bao gồm trình tự mã hoá M protein không hoàn thiện) và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV160+ (CPMVX+, trong đó X=160). NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ; TMCT: đuôi tế bào chất của miền xuyên màng. Fig.37A thể hiện trình tự nucleotit của HA B Wisconsin (PrL-)+H1 California TMCT (SEQ ID NO: 103). Fig.37B thể hiện trình tự axit amin của PDISP/HA B Wisconsin (PrL-)+H1 California TMCT (SEQ ID NO: 104). Fig.37C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1454 (2X35S/CPMV HT; cấu trúc tham khảo). Fig.37D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1893 (2X35S/CPMV160+; cấu trúc được tạo thành chủ yếu từ CPMVX+, trong đó X=160).

Fig.38 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1067 và 1875 (2X35S/CPMV HT PDISP/HA B Brisbane (Prl-)+H1 California TMCT NOS và 2X35S/CPMV HT+ PDISP/HA B Brisbane (PrL-)+H1 California TMCT NOS, một cách tương ứng; xem ví dụ 5.19). Cấu trúc số 1067 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (PDISP/HA B Brisbane (PrL-)+H1 California TMCT). Cấu trúc số 1875 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi bao gồm M protein không hoàn thiện, vị trí tạo đa dòng, và trình tự kozak của thực vật và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV HT+. PDISP: peptit tín hiệu

protein disulfua isomeraza; NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ; TMCT: đuôi tế bào chất của miền xuyên màng. Fig.38A thể hiện trình tự nucleotit của PDISP/HA B Brisbane (PrL-)+H1 California TMCT (SEQ ID NO: 105). Fig.38B thể hiện trình tự axit amin của PDISP/HA B Brisbane (PrL-)+H1 California TMCT (SEQ ID NO: 106). Fig.38C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1875 (2X35S/CPMV160+).

Fig.39 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2072 và 2052 (2X35S/CPMV HT PDISP/HA B Massachusetts (PrL-) NOS và 2X35S/CPMV HT+ PDISP/HA B Massachusetts (PrL-) NOS, một cách tương ứng; xem ví dụ 5.20). Cấu trúc số 2072 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (PDISP/HA B Massachusetts (PrL-)). Cấu trúc số 2052 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi bao gồm M protein không hoàn thiện, vị trí tạo đa dòng, và trình tự kozak của thực vật và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV HT+. PDISP: peptit tín hiệu protein disulfua isomeraza; NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ. Fig.39A thể hiện trình tự nucleotit của PDISP/HA B Massachusetts (PrL-) (SEQ ID NO: 107). Fig.39B thể hiện trình tự axit amin của PDISP/HA B Massachusetts (PrL-) (SEQ ID NO: 108). Fig.39C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2052 (2X35S/CPMV HT+).

Fig.40 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2074 và 2062 (2X35S/CPMV HT PDISP/HA B Massachusetts (PrL-)+H1 California TMCT NOS và 2X35S/CPMV HT+ PDISP/HA B Massachusetts (PrL-)+H1 California TMCT NOS, một cách tương ứng; xem ví dụ 5.21). Cấu trúc số 2074 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (PDISP/HA B Massachusetts (PrL-)+H1 California TMCT). Cấu trúc số 2062 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi bao gồm M protein không hoàn thiện, vị trí tạo đa dòng, và trình tự kozak của

thực vật và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV HT+. PDISP: peptit tín hiệu protein disulfua isomeraza; NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ; TMCT: đuôi tế bào chất của miền xuyên màng. Fig.40A thể hiện trình tự nucleotit của PDISP/HA B Massachusetts (PrL-)+H1 California TMCT (SEQ ID NO: 109). Fig.40B thể hiện trình tự axit amin của PDISP/HA B Massachusetts (PrL-)+H1 California TMCT (SEQ ID NO: 110). Fig.40C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2062 (2X35S/CPMV HT+).

Fig.41 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1445 và 1839 (2X35S/CPMV HT HA B Wisconsin (PrL-) NOS, và 2X35S/CPMV. HT+ HA B Wisconsin (PrL-) NOS, một cách tương ứng; xem ví dụ 5.22). Cấu trúc số 1445 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (HA B Wisconsin (PrL-)). Cấu trúc số 1839 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi bao gồm M protein không hoàn thiện, vị trí tạo đa dòng, và trình tự kozak của thực vật và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV HT+. PrL-: vòng phân giải protein được xóa bỏ; NOS: đoạn kết thúc nopalin synthaza. Fig.41A thể hiện trình tự nucleotit của HA B Wisconsin (PrL-) (SEQ ID NO: 111). Fig.41B thể hiện trình tự axit amin của HA B Wisconsin (PrL-) (SEQ ID NO: 112). Fig.41C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1839 (2X35S/CPMV HT+).

Fig.42 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1454 và 1860 (2X35S/CPMV HT HA B Wisconsin (PrL-)+H1 California TMCT NOS và 2X35S/CPMV HT+ HA B Wisconsin (PrL-)+H1 California TMCT NOS, một cách tương ứng; xem ví dụ 5.23). Cấu trúc số 1454 hợp nhất trình tự CPMV-HT của kỹ thuật hiện nay (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (HA B Wisconsin (PrL-)+H1 California TMCT). Cấu trúc số 1860 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi bao gồm n M protein không hoàn thiện, vị trí tạo đa dòng, và trình tự kozak của thực vật và là một ví dụ về cấu trúc được

tạo thành chủ yếu từ CPMV HT+. NOS: đoạn kết thúc nopalin synthaza; PrL-: vòng phân giải protein được xóa bỏ; TMCT: đuôi tế bào chất của miền xuyên màng. Fig.42A thể hiện trình tự nucleotit của HA B Wisconsin (PrL-)+H1 California TMCT (SEQ ID NO: 113). Fig.42B thể hiện trình tự axit amin của PDISP/HA B Wisconsin (PrL-)+H1 California TMCT (SEQ ID NO: 114). Fig.42C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1893 (2X35S/CPMV HT+).

Fig.43 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 489 (2X35S/CPMV HT H5 Indonesia NOS xem ví dụ 5.24). Cấu trúc số 489 hợp nhất trình tự CPMV-HT (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với trình tự mã hoá M protein không hoàn thiện) và không bao gồm trình tự kozak khác loại giữa 5'UTR và trình tự nucleotit quan tâm (PDISP/H1 California). Fig.43A thể hiện trình tự nucleotit của H5 Indonesia (SEQ ID NO: 115) nguyên thể. Fig.43B thể hiện trình tự axit amin của H5 Indonesia (SEQ ID NO: 116) nguyên thể. Fig.43C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 489 (2X35S/CPMV HT; cấu trúc tham khảo).

Fig.44 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1800 (A-2X35S CPMV160+ PDISP H3Victoria NOS; xem ví dụ 5.25). Cấu trúc số 1800 bao gồm CPMV 5'UTR bao gồm 160 nucleotit, mảnh nhồi (vị trí tạo đa dòng), và trình tự kozak của thực vật (cấu trúc này không bao gồm trình tự mã hoá M protein không hoàn thiện) và là một ví dụ về cấu trúc được tạo thành chủ yếu từ CPMV160+ (CPMVX+, trong đó X=160). PDISP: peptit tín hiệu protein disulfua isomeraza. NOS: đoạn kết thúc nopalin synthaza. Fig.44A thể hiện trình tự đoạn mồi IF**(SacII)-PDI.s1+4c (SEQ ID NO:117). Fig.44B thể hiện trình tự đoạn mồi IF-H3V36111.s1-4r (SEQ ID NO: 118). Fig.44C thể hiện trình tự của PDISP/H3 Victoria (SEQ ID NO:119). Fig.44D thể hiện sự biểu diễn theo sơ đồ của cấu trúc 2171 (vị trí enzym giới hạn SacII và StuI được sử dụng để tuyến tính hoá plasmit được chỉ định). Fig.44E thể hiện cấu trúc 2171 từ bên trái sang bên phải đường biên t-ADN (gạch dưới), 2X35S/CPMV160+/NOS với cátxét biểu hiện chất ức chế Plastocyanine-P19-Plastocyanine, đuôi tế bào chất chuyển qua màng H1 California và CPMV3'UTR (SEQ ID NO: 120). Fig.44F thể hiện cátxét biểu hiện số 1800 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. Trình tự nucleotit được gạch dưới PDISP/H3

Victoria; 5'UTR được thể hiện là nét đậm; trình tự kozak của thực vật gạch dưới đôi; mảnh nhồi (vị trí tạo đa dòng) chứa 16 cặp bazơ được định vị giữa 5'UTR và trình tự kozak của thực vật (SEQ ID NO:121). Fig.44G thể hiện trình tự axit amin của PDISP/H3 Victoria (SEQ ID NO:122). Fig.44H thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 1800 (cấu trúc được tạo thành chủ yếu từ CPMVX+, trong đó X=160).

Fig.45 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 1819 (2X35S CPMV-HT+ PDISP H3Victoria NOS). Cấu trúc số 1819 hợp nhất trình tự CPMV-HT+ (CMPV 5'UTR với codon khởi đầu đột biến ở vị trí 161 được dung hợp với mảnh nhồi mã hoá M protein không hoàn thiện, vị trí tạo đa dòng, và bao gồm trình tự kozak của thực vật giữa vị trí tạo đa dòng và trình tự nucleotit quan tâm (PDISP/H3 Victoria)). PDISP: peptit tín hiệu protein disulfua isomeraza. NOS: đoạn kết thúc nopalin synthaza. Fig.45A thể hiện trình tự đoạn mồi IF(SacII)-Kozac_PDI.c (SEQ ID NO: 123). Fig.45B thể hiện trình tự đoạn mồi IF-H3V36111.s1-4r (SEQ ID NO:124). Fig.45C thể hiện sự biểu diễn theo sơ đồ của cấu trúc 2181. Fig.45D thể hiện trình tự đối với cấu trúc 2181 (từ bên trái sang bên phải đường biên t-ADN, gạch dưới; 2X35S/CPMV-HT+/NOS cátxét biểu hiện chất ức chế Plastocyanine-P19-Plastocyanine; SEQ ID NO:126). Fig.45E thể hiện cátxét biểu hiện số 1819 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. Trình tự nucleotit PDISP/H3 Victoria được gạch dưới (SEQ ID NO:127). Fig.45F thể hiện sự biểu diễn theo sơ đồ của cấu trúc 1819.

Fig.46A thể hiện hoạt tính hemagglutin hoá tương đối của H7 Hangzhou HA nguyên thể và H7 Hangzhou HA được cải biến, với vòng phân giải protein được xóa bỏ khi đồng biểu hiện với M2. H7 Hangzhou HA nguyên thể và được cải biến (cấu trúc #2142 và 2152, xem ví dụ 5.33 và 5.34) được biểu hiện với sự có mặt của M2 (cấu trúc #1261 xem ví dụ 5.1) và được tinh chế từ thực vật. Fig.46B thể hiện sản lượng protein của H7 Hangzhou HA nguyên thể (cấu trúc #2142) và H7 Hangzhou HA được cải biến, với vòng phân giải protein loại bớt (cấu trúc #2152). Fig.46C thể hiện phép phân tích SDS-PAGE, với dãy 2 thể hiện H7 Hangzhou HA được cải biến được tinh chế với vòng phân giải protein được loại bỏ (cấu trúc #2152) và dãy 3 thể hiện H7 Hangzhou HA nguyên thể

được tinh chế (cấu trúc #2142). Đối với mỗi dãy, 2µg tổng protein được nạp vào gel. Việc tinh chế profin protein là tương tự đối với cả hai cấu trúc.

Fig.47A thể hiện sức kháng Trypsin giữa protein HA nguyên thể và HA được cải biến với vòng phân giải protein được thay thế bằng cầu liên kết GG (prl-), HA được cải biến với vòng phân giải protein được thay thế bằng cầu liên kết TETQ (TETQ) và HA được cải biến với vòng phân giải protein được thay thế bằng cầu liên kết TETR (TETR). Các cấu trúc (# 489) nguyên thể, PRL-(# 928), TETQ (#766) và TETR (#676) H5 Indonesia HA VLP được tinh chế. Đối với mỗi dãy, hai mẫu HA VLP được tạo huyền phù lại trong chất đệm (100mM Na/KP_{O4}, 150 mM NaCl, 0,01% TWEEN 80) ở độ pH=7,4 ở nồng độ đích 150µg/mL. Trypsin được bổ sung với tỷ lệ 1:100 vào một mẫu được tạo huyền phù lại. Ủ mẫu trong 30, 60 và 120 phút ở nhiệt độ trong phòng. 30µL phần chiết không tiêu hoá (đối chứng) và 30µL phần chiết tiêu hoá được nạp vào SDS-PAGE gel, mà được nhuộm bằng xanh Coomassie. Fig.47B thể hiện tính gây miễn dịch (độ chuẩn HI) của H5 VLP nguyên thể và bản sao thể đột biến của nó (prl-, TETQ và TETR) ở chuột sau hai liều. Các cột thể hiện phép so sánh độ chuẩn (%) HI tương đối của mỗi thể đột biến H5 VLP với H5 VLP nguyên thể.

Fig.48 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2220 (2X35S/CPMV HT+/ PDISP/H2 Singapore/ NOS xem ví dụ 5.27). Fig.48A thể hiện trình tự nucleotit của đoạn mồi IF**-H2S157.s1-6r (SEQ ID NO: 127). Fig.48B thể hiện trình tự nucleotit của PDISP/H2 Singapore (SEQ ID NO: 128) Fig.48C thể hiện trình tự nucleotit của cátxét biểu hiện số 2220 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. Trình tự nucleotit PDISP/H2 Singapore được gạch dưới (SEQ ID NO: 129). Fig.48D thể hiện trình tự axit amin của PDISP/H2 Singapore (SEQ ID NO: 130). Fig.48E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2220.

Fig.49 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2221 (2X35S/CPMV HT+/ PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ/ NOS xem ví dụ 5.28). Fig.49A thể hiện trình tự nucleotit của đoạn mồi H2S157(Prl-).r (SEQ ID NO: 131). Fig.49B thể hiện trình tự nucleotit của đoạn mồi H2S157(Prl-).c (SEQ ID NO: 132) Fig.49C thể hiện trình tự nucleotit của cátxét biểu hiện số 2221 từ đoạn khở đầu

2X35S đến đoạn kết thúc NOS. Trình tự nucleotit PDISP/H2 Singapore được gạch dưới (SEQ ID NO: 133). Fig.49D thể hiện trình tự axit amin của PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ (SEQ ID NO:134). Fig.49E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2221.

Fig.50 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2222 (2X35S/CPMV 160+/ PDISP/H2 Singapore) và 2223 (2X35S/CPMV 160+/ PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ/ NOS) xem ví dụ 5.29). Fig.50A thể hiện trình tự nucleotit của cátxét biểu hiện số 2222 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. Trình tự nucletotit PDISP/H2 Singapore được gạch dưới (SEQ ID NO:135). Fig.50B thể hiện trình tự nucleotit của cátxét biểu hiện số 2223 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ trình tự nucleotit được gạch dưới (SEQ ID NO: 136). Fig.50C thể hiện theo sơ đồ của cấu trúc số 2222 Fig.50D thể hiện theo sơ đồ của cấu trúc số 2223.

Fig.51 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2219 (2X35S/CPMV HT+ PDISP/H3 Perth) và 2139 (2X35S/CPMV 160+/ PDISP/H3 Perth) xem ví dụ 5.30). Fig.51A thể hiện trình tự nucleotit của PDISP/H3 Perth (SEQ ID NO: 137). Fig.51B thể hiện trình tự nucleotit của đoạn mồi IF**-H3P1609.S1-6r (SEQ ID NO: 138). Fig.51C thể hiện trình tự axit amin của PDISP/H3 Perth (SEQ ID NO: 139). Fig.51D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2219. Fig.51E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2139.

Fig.52 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2039 (2X35S/CPMV HT+ PDISP/H3 Perth với vòng phân giải protein được xóa bỏ) và 2159 (2X35S/CPMV 160+ PDISP/H3 Perth với vòng phân giải protein được xóa bỏ) xem ví dụ 5.31). Fig.52A thể hiện trình tự nucleotit của PDISP/H3 Perth với vòng phân giải protein được xóa bỏ (SEQ ID NO: 140). Fig.52B thể hiện trình tự nucleotit của đoạn mồi H3P1609(Prl-)#2.r (SEQ ID NO: 141). Fig.52C thể hiện trình tự nucleotit của đoạn mồi H3P1609(Prl-)#2.c (SEQ ID NO: 142). Fig.52Dthể hiện trình tự axit amin của PDISP/H3 Perth với vòng phân giải protein được xóa bỏ (SEQ ID NO: 143). Fig.52E thể hiện phép

biểu diễn theo sơ đồ của cấu trúc số 2039. Fig.52F thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2159.

Fig.53 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2230 (2X35S/CPMV HT+ PDISP/H3 Victoria với vòng phân giải protein được xóa bỏ) và 2250 (2X35S/CPMV 160+ PDISP/H3 Victoria với vòng phân giải protein được xóa bỏ) xem ví dụ 5.32). Fig.53A thể hiện trình tự nucleotit của trình tự nucleotit của PDISP/H3 Victoria với vòng phân giải protein được xóa bỏ (SEQ ID NO: 144). Fig.53B thể hiện trình tự nucleotit của đoạn mồi H3V36111(Prl-).r (SEQ ID NO: 145). Fig.53C thể hiện trình tự nucleotit của đoạn mồi H3V36111(Prl-).c (SEQ ID NO: 146). Fig.53D thể hiện trình tự axit amin của PDISP/H3 Victoria với vòng phân giải protein được xóa bỏ (SEQ ID NO: 147). Fig.53E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2230. Fig.53F thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2250.

Fig.54 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2142 (2X35S/CPMV HT+/PDISP/H7 Hangzhou/NOS) xem ví dụ 5.33). Fig.54A thể hiện trình tự nucleotit của PDISP/H7 Hangzhou (SEQ ID NO: 148). Fig.54B thể hiện trình tự nucleotit của đoạn mồi IF*-H7H113.s1-6r (SEQ ID NO: 149). Fig.54C thể hiện trình tự axit amin của PDISP/H7 Hangzhou (SEQ ID NO: 150). Fig.54D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2142.

Fig.55 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2152 (2X35S/CPMV HT+/PDISP/H7 Hangzhou với vòng phân giải protein được xóa bỏ/NOS) xem ví dụ 5.34). Fig.55A thể hiện trình tự nucleotit của PDISP/H7 Hangzhou với vòng phân giải protein được xóa bỏ (SEQ ID NO: 151). Fig.55B thể hiện trình tự nucleotit của đoạn mồi H7H113(PrL-).r (SEQ ID NO: 152). Fig.55C thể hiện trình tự nucleotit của đoạn mồi H7H113(PrL-).c (SEQ ID NO: 153). Fig.55D thể hiện trình tự axit amin của PDISP/H7 Hangzhou với vòng phân giải protein được xóa bỏ (SEQ ID NO: 154). Fig.53E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2152.

Fig.56 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2224 (2X35S/CPMV HT+ PDISP/H9 Hong Kong) và 2226 (2X35S/CPMV 160+ PDISP/H9

Hong Kong) xem ví dụ 5.35). Fig.56A thể hiện trình tự nucleotit của PDISP/H9 Hong Kong (SEQ ID NO: 155). Fig.56B thể hiện trình tự nucleotit của đoạn mồi IF**-H9HK107399.S1-6r (SEQ ID NO: 156). Fig.56C thể hiện trình tự axit amin của PDISP/H9 Hong Kong (SEQ ID NO: 157). Fig.56D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2224. Fig.56E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2226.

Fig.57 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2225 (2X35S/CPMV HT+ PDISP/H9 Hong Kong với vòng phân giải protein được xóa bỏ) và 2227 (2X35S/CPMV 160+ PDISP/H9 Hong Kong với vòng phân giải protein được xóa bỏ) xem ví dụ 5.36. Fig.57A thể hiện trình tự nucleotit của PDISP/H9 Hong Kong với vòng phân giải protein được xóa bỏ (SEQ ID NO: 158). Fig.57B thể hiện trình tự nucleotit của đoạn mồi H9HK107399(Prl-).r (SEQ ID NO: 159). Fig.57C thể hiện trình tự nucleotit của đoạn mồi H9HK107399(Prl-).c (SEQ ID NO: 160). Fig.57D thể hiện trình tự axit amin của PDISP/H9 Hong Kong với vòng phân giải protein được xóa bỏ (SEQ ID NO: 161). Fig.57E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2225. Fig.57F thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2227.

Fig.58 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2013 (2X35S/CPMV 160+/PDISP/HA B Malaysia/NOS) xem ví dụ 5.37. Fig.58A thể hiện trình tự nucleotit của PDISP/HA B Malaysia (SEQ ID NO: 162). Fig.58B thể hiện trình tự nucleotit của đoạn mồi IF**-HBM250604.S1-6r (SEQ ID NO: 163). Fig.58C thể hiện trình tự axit amin của PDISP/HA B Malaysia (SEQ ID NO: 164). Fig.58D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2013.

Fig.59 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2014 (2X35S/CPMV 160+/PDISP/HA B Malaysia với vòng phân giải protein được xóa bỏ/NOS) xem ví dụ 5.38. Fig.59A thể hiện trình tự nucleotit của PDISP/HA B Malaysia với vòng phân giải protein được xóa bỏ (SEQ ID NO: 165). Fig.59B thể hiện trình tự nucleotit của đoạn mồi HBM250604(PrL-).r (SEQ ID NO: 166). Fig.59C thể hiện trình tự nucleotit của đoạn mồi HBM250604(PrL-).c (SEQ ID NO: 167). Fig.59D thể hiện trình tự axit amin của PDISP/HA B Malaysia với vòng phân giải protein được xóa bỏ (SEQ ID NO: 168). Fig.59E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2014.

Fig.60 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2070 (2X35S/CPMV HT PDISP/HA B Massachusetts), 2080 (2X35S/CPMV HT+ PDISP/HA B Massachusetts) và 2090 (2X35S/CPMV 160+ PDISP/HA B Massachusetts) xem ví dụ 5.39. Fig.60A thể hiện trình tự nucleotit của PDISP/HA B Massachusetts (SEQ ID NO: 169). Fig.60B thể hiện trình tự axit amin của PDISP/HA B Massachusetts (SEQ ID NO: 170). Fig.60C thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2070. Fig.60D thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2080. Fig.60E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2090.

Fig.61 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2102 (2X35S/CPMV HT PDISP/HA B Florida với vòng phân giải protein được xóa bỏ) và 2104 (2X35S/CPMV HT+ /BeYDV/PDISP/HA B Florida với vòng phân giải protein được xóa bỏ) xem ví dụ 5.40. Fig.61A thể hiện trình tự nucleotit của đoạn mồi HBF406(PrL-).r (SEQ ID NO: 190). Fig.61B thể hiện trình tự nucleotit của đoạn mồi HBF406(PrL-).c (SEQ ID NO: 191). Fig.61C thể hiện trình tự nucleotit của đoạn mồi IF*-HBF406.s1-6r (SEQ ID NO: 192). Fig.61D thể hiện trình tự nucleotit của PDISP/HA B Florida với vòng phân giải protein được xóa bỏ (SEQ ID NO: 193). Fig.61E thể hiện trình tự axit amin của PDISP/HA B Florida với vòng phân giải protein được xóa bỏ (SEQ ID NO: 194). Fig.61F thể hiện trình tự nucleotit của cátxét biểu hiện số 2102 (SEQ ID NO: 195). Fig.61G thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2102. Fig.61H thể hiện trình tự nucleotit của cátxét biểu hiện số 2104 (SEQ ID NO: 196). Fig.61I thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2104.

Fig.62 thể hiện thành phần trình tự được sử dụng để tạo ra cấu trúc số 2106 (2X35S/CPMV HT+/ PDISP/B Florida +H1 California TMCT với vòng phân giải protein được xóa bỏ/NOS) và 2108 (2X35S/CPMV HT+/ BeYDV/PDISP/B Florida +H1 California TMCT với vòng phân giải protein được xóa bỏ/ NOS) xem ví dụ 5.41. Fig.62A thể hiện trình tự nucleotit của đoạn mồi IF-H1cTMCT.S1-4r (SEQ ID NO: 197). Fig.62B thể hiện trình tự nucleotit của PDISP/HA B Florida+H1Cal TMCT với vòng phân giải protein được xóa bỏ (SEQ ID NO: 198). Fig.62C thể hiện trình tự axit amin của PDISP/HA B Florida+H1Cal TMCT với vòng phân giải protein được xóa bỏ (SEQ ID

NO: 199). Fig.62D thể hiện trình tự nucleotit của cátxét biểu hiện số 2106 (SEQ ID NO: 200). Fig.62E thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2106. Fig.62F thể hiện trình tự nucleotit của cátxét biểu hiện số 2108. Fig.62G thể hiện phép biểu diễn theo sơ đồ của cấu trúc số 2108.

Mô tả chi tiết sáng chế

Phần mô tả sau đây là phương án được ưu tiên.

Sáng chế đề cập đến hạt tương tự virut (VLP) và phương pháp tạo ra và làm tăng sản lượng, sự tích tụ và sản xuất VLP ở thực vật.

Sáng chế đề xuất, một phần, phương pháp tạo ra hạt tương tự virut (VLP) ở thực vật, hoặc phần của thực vật. Phương pháp này bao gồm việc đưa axit nucleic vào trong thực vật hoặc phần của thực vật. Axit nucleic bao gồm bao gồm vùng điều hoà có hoạt tính ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá hemagglutinin (HA) cúm. HA bao gồm vòng phân giải protein được cải biến hoặc vị trí tách. Thực vật hoặc phần của thực vật được ủ trong các điều kiện mà cho phép sự biểu hiện của axit nucleic, do đó tạo ra VLP. Nếu muốn, thực vật hoặc phần của thực vật có thể được thu hoạch và VLP được tinh chế.

Sáng chế cũng đề xuất VLP được tạo ra bằng phương pháp này. VLP có thể bao gồm một hoặc nhiều hơn một lipit thu được từ thực vật.

VLP có thể được sử dụng để tạo ra chế phẩm bao gồm liều hữu hiệu của VLP để gây ra đáp ứng miễn dịch, và chất mang dược dụng.

Cũng được đề xuất trong bản mô tả này là hemagglutinin được cải biến, trong đó vòng phân giải protein hoặc vị trí tách đã được cải biến.

Sáng chế cũng đề xuất vật liệu thực vật bao gồm VLP được tạo ra bằng cách biểu hiện axit nucleic được mô tả trên đây. Vật liệu thực vật có thể được sử dụng trong việc gây miễn dịch đối với sự lây nhiễm virut cúm ở đối tượng. Vật liệu thực vật cũng được trộn như thành phần bổ sung thức ăn.

VLP của sáng chế cũng có thể được tạo ra bằng việc cung cấp thực vật hoặc phần của thực vật bao gồm axit nucleic như được xác định trên đây và ủ thực vật hoặc phần của thực vật trong các điều kiện mà cho phép sự biểu hiện của axit nucleic, do đó tạo ra VLP. VLP có thể bao gồm một hoặc nhiều hơn một lipit thu được từ thực vật. VLP có thể được sử dụng để tạo ra chế phẩm bao gồm liều hữu hiệu của VLP để gây ra đáp ứng miễn dịch, và chất mang dược dụng. Sáng chế cũng đề xuất vật liệu thực vật bao gồm VLP được tạo ra bằng cách biểu hiện axit nucleic thứ nhất và thứ hai. Vật liệu thực vật có thể được sử dụng để gây miễn dịch đối với sự lây nhiễm virut cúm ở đối tượng. Vật liệu thực vật cũng có thể được trộn như thành phần bổ sung cho thực phẩm.

VLP của sáng chế bao gồm một hoặc nhiều hemagglutinin (HA) cúm được cải biến. HA được cải biến có thể thu được từ HA bất kỳ, ví dụ H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16 hoặc typ B HA như được mô tả trong các công bố đơn quốc tế số WO 2009/009876; WO 2009/076778; WO 2010/003225; WO 2010/003235; WO 2011/03522).

Sáng chế bao gồm VLP bao gồm trình tự HA của chủng cúm, trong đó trình tự HA bao gồm vị trí tách polybazơ được cải biến bao gồm ví dụ, các cải biến như được mô tả trong bản mô tả này.

HA Protein

Thuật ngữ "hemagglutinin" hoặc "HA" như được sử dụng trong bản mô tả này đề cập đến glycoprotein được tìm thấy ở phía bên ngoài của hạt virut cúm. HA là typ I glycoprotein chuyển qua màng homotrime, thường bao gồm peptit tín hiệu, miền HA1 và miền HA2 bao gồm vị trí neo mở rộng màng ở đầu cuối C và đuôi tế bào chất nhỏ. Các trình tự nucleotit mã hoá HA cũng đã được biết đến và sẵn có xem, ví dụ ấn phẩm: BioDefence Public Health base (virut cúm; xem URL: biohealthbase.org) hoặc National Center for Biotechnology Information (xem URL: ncbi.nlm.nih.gov). HA có thể bao gồm HA bất kỳ, ví dụ H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16 hoặc typ B HA như được mô tả trong công bố đơn quốc tế số WO 2009/009876; WO 2009/076778; WO 2010/003225; WO 2010/003235; WO 2011/03522). Hơn nữa, HA có

thể được tạo thành từ trình tự của hemagglutinin mà được tách từ một hoặc nhiều virut cúm mới xuất hiện hoặc được nhận dạng là mới. Sáng chế cũng bao gồm VLP mà bao gồm các HA được cải biến thu được từ một hoặc nhiều hơn một typ phụ gây cúm.

HA monome có thể được chia thành ba miền chính – miền thân hoặc cụm miền thân (SDC), miền đầu hình cầu hoặc cụm miền đầu (HDC) và cụm miền xuyên màng (TDC). SDC và HDC có thể được đề cập chung đến như 'ectodomain'. Tài liệu công bố bởi Ha et al. 2002 (EMBO J. 21 :865-875) mô tả sự định hướng tương đối của các miền phụ khác nhau của SDC và HDC trong vài kiểu phụ gây cúm, dựa vào các cấu trúc tinh thể học tia X.

HA protein được tổng hợp như tiền chất protein (HA0) có khoảng 75kDa, mà có thể lắp ráp ở bề mặt thành protein trime kéo dài. Tiền chất protein được tách ở vị trí tách hoạt hoá bảo tồn thành 2 chuỗi polypeptit, HA1 và HA2 (bao gồm vùng chuyển qua màng), được liên kết bởi liên kết disulfua. Fig.15 thể hiện ví dụ không hạn chế về các trình tự axit amin của vùng liên kết đối với vài HA.

Thuật ngữ "homotrime" hoặc "homotrime" cho biết rằng oligome được tạo thành bởi ba phân tử protein HA. Không muốn bị giới hạn bởi lý thuyết, protein HA được tổng hợp như tiền chất protein (HA0) monome khoảng 75kDa, mà có thể lắp ráp ở bề mặt thành protein trime kéo dài. Trước khi sự trime hoá xảy ra, tiền chất protein được tách ở vị trí tách hoạt hoá bảo tồn (cũng đề cập đến peptit dung hợp) thành 2 chuỗi polypeptit, HA1 và HA2 (bao gồm vùng chuyển qua màng), được liên kết bởi liên kết disulfua. Các đoạn HA1 có thể có chiều dài 328 axit amin và đoạn HA2 có thể có chiều dài 221 axit amin. Mặc dù việc tách này có thể là quan trọng đối với khả năng lây nhiễm virut, có thể không nhất thiết đối với sự trime hoá của protein hoặc đối với khả năng gây miễn dịch. Việc xen vào của HA nằm trong màng mạng lưới nội chất (ER) của tế bào chủ, việc tách peptit tín hiệu và glycosyl hoá protein là sự kiện đồng dịch mã. Việc gấp nếp chính xác của HA đòi hỏi sự glycosyl hoá của protein và sự hình thành từ 5 đến 6 liên kết disulfua liên chuỗi. HA trime có thể lắp ráp trong phức hệ cis- và trans-Golgi, miền xuyên màng đóng vai trò trong quy trình trime hoá. Cấu trúc tinh thể của protein HA được xử lý bromelain, mà thiếu miền xuyên màng, đã thể hiện cấu trúc bảo tồn ở mức cao trong số các chủng cúm. Cũng đã

được thiết lập rằng HA trải qua các thay đổi hình dạng lớn trong quy trình lây nhiễm, mà đòi hỏi tiền chất HA0 để được tách thành 2 chuỗi polypeptit HA1 và HA2. protein HA có thể được xử lý (nghĩa là, bao gồm các miền HA1 và HA2) hoặc có thể không được xử lý (nghĩa là bao gồm miền HA0). Tiền chất protein không được xử lý của HA được tổng hợp như tiền chất protein (HA0) có khoảng 75kDa, mà có thể lắp ráp ở bề mặt thành protein trime kéo dài. Tiền chất protein được tách ở vị trí tách bảo tồn (cũng được biết là vòng phân giải protein) thành 2 chuỗi polypeptit, HA1 và HA2 (bao gồm vùng chuyển qua màng), được liên kết bằng liên kết disulfua.

HA protein như được mô tả trong bản mô tả này còn có thể là protein HA được cải biến (cũng được đề cập đến như "thể đột biến HA"), ví dụ tiền chất protein (HA0) được cải biến, trong đó vòng phân giải protein hoặc vị trí tách được cải biến.

HA/vị trí tách được cải biến

Sau khi tách HA0, HA trở nên nhạy với độ pH, trải qua sự thay đổi hình dạng không thuận nghịch ở độ pH của nội thể (<pH=6,0). Hình dạng của tiền chất HA0 là ổn định ở độ pH thấp, nhưng dạng HA1-HA2 được tách, là nửa ổn định (Bullough PA et. al., 1994, Nature. Vol 371 :37-43). Ngưỡng độ pH mà gây ra sự thay đổi hình dạng trong các HA khác nhau là khoảng độ pH từ 5,8 đến 5,9 đối với các chủng B, trong khi nó có tính axit nhiều hơn, độ pH từ 5,1 đến 5,8, đối với typ A HA (Beyer WEP et al, 1986, Archives Virol, vol 90: 173). Sau khi tách, đầu cuối amino của HA2 là trình tự không có cực chứa 23 axit amin mà sau đó trở thành miền xuyên màng bong ra ngang qua màng sinh chất vật chủ (được gọi là peptit dung hợp; Fig.15). Vị trí tách của HA được định vị ở vòng nhô mà bề mặt của HA, và vị trí này có thể đạt đến bởi proteaza.

Để tối ưu hoá việc sản xuất vacxin trong trứng và duy trì hoạt tính nhưng virut giảm độc lực, việc cải biến vị trí tách polybazơ của H5 (RERRRKKR↓G) đã được nghiên cứu (Horimoto T, et. al, 2006, Vaccine, Vol 24 : 3669-3676). Thể đột biến quan tâm chứa sự xóa bỏ 4 axit amin mang điện tích thứ nhất (RERR) và sự thay thế axit amin RKKR bằng TETR mà làm bất hoạt vị trí tách polybazơ nhưng duy trì khả năng xử lý HA0 thành HA1-HA2 qua gốc Arginin của TETR motif (xem Fig.19). Chiến lược tương tự để tạo ra

virut giảm độc lực được dùng bởi NIBSC để huỷ bỏ vị trí polybazơ cho phép tạo ra ở sản lượng cao chủng A/Turkey/Turkey/1/2005 H5N1 mà không giết trứng. Trình tự vị trí polybazơ (GERRRKKR↓G) được thay thế bằng RETR trong thể độ biến của chúng (NIBSC 05/240 virut tham khảo cúm NIBSC NIBG-23). Vị trí tách polybazơ của H5 HA cũng đã được thay thế bằng vị trí monobazơ của H6 đối với sự biểu hiện trong trứng. Trong ví dụ này, 4 gốc thứ nhất và bốn gốc axit amin sau cùng của vị trí polybazơ được thay thế bằng IETR (sự thay thế của RERRRKKR↓G bằng IETR↓G; Hoffman E, et. al., 2002, Vaccine, Vol 20 :3165-3170). Trong mỗi trong số các ví dụ được đề xuất trên đây, sự cải biến được thực hiện để làm giảm độc lực virut trong khi duy trì việc tạo ra HA trong trứng. Tức là, việc tách tiền chất HA0 là không được làm bất hoạt một cách tổng thể để cho phép HA0 được xử lý thành HA1-HA2 và trải qua sự thay đổi hình dạng theo độ pH, do đó cho phép sao chép virut trong tế bào vật chủ.

Như được sử dụng trong bản mô tả này, thuật ngữ "hemagglutinin được cải biến" hoặc "HA được cải biến", "hemagglutinin đột biến" hoặc “HA đột biến” đề cập đến HA trong đó HA có sự cải biến hoặc đột biến, ví dụ sự thay thế, xen vào, loại bớt hoặc tổ hợp của chúng, mà dẫn đến trình tự axit amin được thay đổi trong vòng phân giải protein hoặc vị trí tách của protein HA.

Cấu trúc tinh thể của HA0 từ A/Hong Kong/68 đã được xác định (Chen, J., 1998. Cell 95:409-417). Các gốc mà được phơi trần với dung môi thường được cho là một phần của vị trí tách mà tạo thành vòng bề mặt kéo dài được phơi trần ở mức cao. Trình tự liên ứng có thể được xác định trong vùng được chọn này chẳng hạn, nhưng không chỉ giới hạn ở:

A/H3/HA0 liên ứng: NVPEKQTR/GIFGAIAGFIE (SEQ ID NO: 66)

A/H1/HA0 liên ứng: NIPSIQSR/GLFGAIAGFIE (SEQ ID NO: 67)

H5 của loại chim liên ứng: QRESRRKKR/GLFGAIAGFIEG (SEQ ID NO: 1)

B/HA0 liên ứng: PAKLLKER/GFFGAIAGFLE (SEQ ID NO: 68)

Trong đó, việc tách giữa HA1 và HA2 được chỉ định bởi "/" (xem ấn phẩm: Bianchi et al., 2005, Journal of Virology, 79:7380-7388), và cả các Fig.15 và Fig.18A.

HA protein có thể là hemagglutinin cúm typ B hoặc hemagglutinin protein cúm typ A với sự cải biến trong vùng vòng phân giải protein, ví dụ loại bớt, xen vào, thay thế hoặc tổ hợp của chúng của vòng phân giải protein (vị trí tách). Không bị ràng buộc bởi lý thuyết, sự cải biến của vòng phân giải protein có thể bảo đảm rằng phân tử HA được duy trì như tiền chất HA0. Do đó tạo ra VLP tương đồng và kiên định hơn bao gồm HA0 protein.

Thuật ngữ "vòng phân giải protein" hoặc "vị trí tách" có nghĩa là trình tự liên ứng của vị trí phân giải protein mà có liên quan đến việc tách tiền chất HA0. "Liên ứng" hoặc "trình tự liên ứng" như được sử dụng trong bản mô tả này có nghĩa là trình tự (trình tự axit amin hoặc nucleotit) mà bao gồm khả năng thay đổi của trình tự của trình tự liên quan dựa trên phép phân tích căn thẳng hàng của nhiều trình tự, chẳng hạn, kiểu phụ của trình tự HA0 cúm cụ thể. Trình tự liên ứng của vị trí tách HA0 cúm có thể bao gồm trình tự axit amin hemagglutinin liên ứng cúm A, bao gồm ví dụ H1 liên ứng (SEQ ID NO: 67), H3 liên ứng (SEQ ID NO: 66), H5 liên ứng (SEQ ID NO: 1) hoặc trình tự axit amin hemagglutinin liên ứng cúm B (SEQ ID NO: 68), ví dụ nhưng không chỉ giới hạn ở B Florida và B Malaysia. Các ví dụ không hạn chế về trình tự của vùng vòng phân giải protein được thể hiện trên Fig.15 và Fig.18B (và xem ấn phẩm: Bianchi et al., 2005, Journal of Virology, 79:7380-7388).

Các gốc trong vòng phân giải protein hoặc vị trí tách có thể được đột biến, ví dụ nhưng không chỉ giới hạn ở đột biến điểm, thay thế, xen vào hoặc loại bớt. Thuật ngữ "đột biến axit amin" hoặc "cải biến axit amin" như được sử dụng trong bản mô tả này có nghĩa là bao gồm sự thay thế, loại bớt, xen vào axit amin và cải biến. Tổ hợp bất kỳ của sự thay thế, loại bớt, xen vào có thể được tạo ra để đạt được ở cấu trúc cuối cùng, với điều kiện rằng cấu trúc cuối cùng có các đặc điểm mong muốn, ví dụ việc phân cắt hoặc hủy bỏ giảm của vòng phân giải protein hoặc vị trí phân cắt bởi proteaza.

Thuật ngữ "vòng phân giải protein được cải biến" có nghĩa là vòng phân giải protein có thể bao gồm một hoặc nhiều đột biến điểm, xóa bỏ một phần, xóa bỏ hoàn toàn, thay thế một phần bằng trình tự liên kết, thay thế hoàn toàn bằng trình tự liên kết, bao gồm sự thay thế một phần hoặc hoàn toàn axit amin nằm trong vị trí tách bằng một hoặc nhiều axit amin không protein hoặc tổ hợp của chúng. Tương tự, thuật ngữ "vị trí tách được cải biến", có nghĩa là vị trí tách nằm trong vòng phân giải protein có thể bao gồm một hoặc nhiều đột biến điểm, xóa bỏ một phần, xóa bỏ hoàn toàn, thay thế một phần bằng trình tự liên kết, thay thế hoàn toàn bằng trình tự liên kết, bao gồm sự thay thế một phần hoặc hoàn toàn axit amin nằm trong vị trí tách bằng một hoặc nhiều axit amin không protein hoặc tổ hợp của chúng. Các cải biến đối với vòng phân giải protein, vị trí tách, hoặc cả hai, cũng có thể bao gồm việc xóa bỏ, thay thế hoặc thế của một hoặc nhiều axit amin được định vị ở bên ngoài của hoặc liền kề với, vòng phân giải protein hoặc trình tự vị trí tách. Thuật ngữ "cầu liên kết" có nghĩa là trình tự axit amin bao gồm một hoặc nhiều axit amin mà có thể được đưa vào trong vòng phân giải protein hoặc vị trí tách hoặc mà có thể thay thế một vài hoặc tất cả axit amin bằng vòng phân giải protein hoặc vị trí tách. Cầu liên kết có thể được thiết kế để bảo đảm rằng việc xóa bỏ axit amin bất kỳ nằm trong vòng phân giải protein hoặc vị trí tách không phá vỡ sự biểu hiện hoặc hoạt tính tiếp theo của HA được cải biến.

Bằng cách làm ổn định protein HA bằng cách cải biến hoặc xóa bỏ vòng phân giải protein, sản lượng sản phẩm hoặc protein gia tăng là có thể đạt được, khi biểu hiện HA được cải biến ở thực vật, khi so với HA nguyên thể được biểu hiện ở thực vật trong cùng điều kiện. Hơn nữa, bằng cách cải biến hoặc xóa bỏ vòng phân giải protein, khả năng thay đổi sự biểu hiện của HA được cải biến được biểu hiện là được giảm và tính kiên định của HA được cải biến được tạo ra là được gia tăng, khi so với HA nguyên thể được biểu hiện ở thực vật trong cùng điều kiện.

Do đó, sáng chế cũng bao gồm phương pháp làm tăng sản lượng sản phẩm của protein HA ở thực vật. Không muốn bị ràng buộc bởi lý thuyết, các tác giả sáng chế đã cho rằng bằng cách cải biến hoặc xóa bỏ vòng phân giải protein trong protein HA, độ ổn định được cải thiện kháng lại sự thoái hoá phân giải protein ở thực vật, sự ổn định hoá

trong quá trình đi qua của HA trong quy trình tiết của bộ máy golgi, và trong quá trình tinh chế là được tạo ra.

Hơn nữa, sáng chế cũng bao gồm phương pháp làm tăng chất lượng sản phẩm của protein HA biểu hiện ở thực vật. Thuật ngữ chất lượng sản phẩm, có nghĩa là ví dụ sản lượng sản phẩm gia tăng của HA biểu hiện ở thực vật, độ ổn định của sản phẩm ví dụ độ ổn định gia tăng của HA biểu hiện ở thực vật, tính kiên định của sản phẩm ví dụ việc tạo ra sản phẩm đồng nhất ví dụ HA0 hoặc tổ hợp của chúng.

Thuật ngữ về sự gia tăng sản lượng của sản phẩm hoặc protein, có nghĩa là sự gia tăng về sản lượng tương đối protein bằng khoảng từ 20% đến 100%, hoặc lượng bất kỳ giữa chúng như được xác định bằng cách sử dụng các kỹ thuật tiêu chuẩn trong lĩnh vực này, ví dụ, khoảng từ 40% đến 70% hoặc trị số bất kỳ giữa chúng ví dụ khoảng 20, 22, 24, 25, 26, 28, 30, 32, 34, 35, 36, 38, 40, 42, 44, 45, 46, 48, 50, 52, 54, 55, 56, 58, 60, 65, 70, 75, 80, 85, 90, 95, hoặc 100%, hoặc lượng bất kỳ giữa chúng, khi so với sản lượng sản phẩm hoặc protein của cùng protein HA mà không có vòng phân giải protein của nó được loại bỏ.

Như được thể hiện trên các hình vẽ Fig.13A và Fig.14, HA từ B/Brisbane/60/2008 biểu hiện kém ở lá *Nicotiana benthamiana* thấm lọc nông nghiệp (xem dãy 1008). Tuy nhiên, sự biểu hiện của HA typ B mà đã được cải biến để xóa bỏxóa bỏ vòng phân giải protein (xem dãy 1059, Fig.13A, Fig.14) dẫn đến sự biểu hiện gia tăng. Hơn nữa, sự đồng biểu hiện của HA-typ B với M2 từ A/New Caledonia/20/99, dẫn đến gia tăng sự biểu hiện HA (xem các dãy "1008+1261"; và 1059+1261"). Sự đồng biểu hiện của HA typ B bao gồm việc xóa bỏ trong vòng phân giải protein, với M2 từ A/Puerto Rico/8/34 cũng dẫn đến sự biểu hiện gia tăng (1059+859; Fig.14).

Như được thể hiện hơn nữa trên Fig.46B, sản lượng protein của protein HA từ H7 A/Hangzhou/1/13 biểu hiện ở *Nicotiana benthamiana* thấm lọc nông nghiệp được gia tăng ở HA mà đã được cải biến để xóa bỏ hoặc cải biến vòng phân giải protein. Sự đồng biểu hiện của HA nguyên thể (kiểu hoang dại) H7 A/Hangzhou/1/13 với M2 từ A/New Caledonia/20/99 dẫn đến sản lượng protein tương đối 100%, trong khi sự đồng biểu hiện

của HA H7 A/Hangzhou/1/13 bao gồm việc xóa bỏ trong vòng phân giải protein, với M2 từ A/New Caledonia/20/99, dẫn đến sản lượng protein tương đối 182% (Fig.46B). Sự gia tăng sản lượng protein tương đối của HA bao gồm việc xóa bỏ trong vòng phân giải protein là, tuy nhiên, không phụ thuộc vào M2. Ví dụ, như được thể hiện trên Fig.29A, H7 A/Hangzhou/1/13 bao gồm việc xóa bỏ trong vòng phân giải protein thể hiện sự biểu hiện gia tăng (như đo bằng độ chuẩn HA tương đối) khi so với H7 nguyên thể A/Hangzhou/1/13.

Vài chiến lược được đánh giá để làm bất hoạt việc tách HA0 đối với các chủng A và B. Trình tự liên ứng mà được nhận biết bởi proteaza được bao quanh trên vòng mở rộng, phơi trần với dung môi và gần với phần xa màng với protein. Trong chủng B, vòng này chứa 2 motif trình tự được nhận biết bởi proteaza và axit amin đầu cuối N thứ nhất của miền HA2. Phương pháp đột biến điểm (ví dụ xem bảng 2, dưới đây) làm bất hoạt việc tách tiền chất HA0 dẫn đến việc tạo ra HA0, mà không gia tăng sự kết tụ của chủng B VLP. Việc xóa bỏ motif trình tự bao gồm 2 motif tách proteaza (7 axit amin) huỷ bỏ sự kết tụ của B HA. Việc loại bỏ toàn bộ vòng 18 axit amin ra khỏi protein HA của chủng B và xen cầu liên kết vào các dấu hiệu cấu trúc bảo tồn (sợi beta) của cấu trúc protein là hữu hiệu (xem dưới đây; Fig.13A, Fig.14, Fig.16A, Fig.17B). Việc loại bỏ hoặc thay thế của vòng phân giải protein trong protein HA của chủng A là cũng hữu hiệu (xem Fig.20, Fig.22).

Việc xóa bỏ và xen vào trình tự axit amin bao gồm việc xóa bỏ axit amin và xen vào axit amin. Ví dụ không hạn chế về việc xóa bỏ ở cúm B là sự xóa bỏ 17 axit amin (AKLLKERGFFGAIAGFLE) (ví dụ các vị trí axit amin 13-30 của SEQ ID NO: 36) ra khỏi vị trí từ 340 đến 357 của protein HA chín ví dụ như được thể hiện trên Fig.18C đối với cúm B Florida (SEQ ID NO: 72) và B Malaysia (SEQ ID NO: 73). Việc xóa bỏ này có thể được thay thế bằng cầu liên kết thích hợp để liên kết các chuỗi polypeptit để biểu hiện đúng đắn, ví dụ nhưng không chỉ giới hạn ở, sử dụng trình tự "GG", như được thể hiện trên **Fig**.21B (SEQ ID NO:17; B/Brisbane/60/2008 được cải biến; thay thế AKLLKERGFFGAIAGFLEG bằng GG; ví dụ cấu trúc 1059, Fig.5D, Fig.10; Cấu trúc 1039, Fig.8B, hoặc Cấu trúc 1467; Fig.7D, 7E). Việc thay thế khác tiến hành bao gồm việc

thay thế "PPAKLLKER" (ví dụ các vị trí axit amin 339-347 của SEQ ID NO: 16) bằng "GSSSGSSSG", như được thể hiện trên Fig.21C (SEQ ID NO: 18). Hơn nữa, trình tự "RESRRKKR" (ví dụ các vị trí axit amin 339-346 của SEQ ID NO: 116) có thể được thay thế bằng "TETR" (SEQ ID NO: 77) hoặc "TETQ" (SEQ ID NO: 81), như được thể hiện trên Fig.19 đối với cúm H5/Indonesia.

Các đột biến axit amin thay thế đối với HA từ chủng A bao gồm các thay thế axit amin, xen vào và loại bớt, ví dụ nhưng không chỉ giới hạn ở việc xóa bỏ trong vùng vòng phân giải protein của H5 Anhui của trình tự axit amin "RERRRKRGLFGAIAGFIE" (các vị trí axit amin 4-21 của SEQ ID NO: 28), loại bớt trình tự axit amin của vùng vòng phân giải protein của H5 Indo bao gồm "RESRRKKRGLFGAIAGFIE" (các vị trí axit amin 4-22 của SEQ ID NO: 29) hoặc loại bớt trình tự axit amin của vùng vòng phân giải protein của H5 Vietnam "RERRRKKRGLFGAIAGFIE" (các vị trí axit amin 4-22 của SEQ ID NO: 30). Đối với H3, trình tự "RNVPEKQTRGIF" (ví dụ các vị trí axit amin 345-356 của SEQ ID NO: 139) có thể được loại bớt và được thay thế bằng trình tự liên kết thích hợp, ví dụ nhưng không chỉ giới hạn ở "GS" như được thể hiện trên Fig.21E (SEQ ID NO:20). Theo cách khác, trình tự "RNVPEKQTR" (ví dụ các vị trí axit amin 345-353 của SEQ ID NO: 139) trong H3 có thể được thay thế bằng "GSSGSSGSS" như được thể hiện trên Fig.21F (SEQ ID NO: 21; H3 được cải biến A/Perth/16/2009).

Hơn nữa, việc cải biến hoặc làm thay đổi vòng phân giải protein hoặc vị trí tách của HA làm giảm hoặc huỷ bỏ việc tách vòng phân giải protein hoặc vị trí phân cắt bởi proteaza, cũng có thể bao gồm sự thay thế axit amin không bảo tồn, nghĩa là thay thế một axit amin bằng axit amin khác có các đặc tính cấu trúc và/hoặc hoá học khác nhau. Các axit amin không protein cũng có thể được sử dụng cho việc thay thế. Ví dụ, sự thay thế axit amin có thể bao gồm việc thay thế axit amin kỵ nước bằng ưa nước. Sự thay thế axit amin có thể bao gồm sự thay thế bằng axit amin không xuất hiện tự nhiện hoặc bằng dẫn xuất axit amin xuất hiện tự nhiên của axit amin protein.

Các đột biến axit amin đối với HA từ chủng B và/hoặc chủng A có thể bao gồm việc xóa bỏ axit amin. Ví dụ để làm giảm hoặc huỷ bỏ việc tách vòng phân giải protein hoặc vị trí phân cắt bởi proteaza, một hoặc nhiều axit amin là loại bớt hoặc loại bỏ nằm

trong vòng phân giải protein hoặc trình tự vị trí tách. Các ví dụ không hạn chế về việc xóa bỏ bao gồm việc loại bỏ axit amin từ 323 đến 341 của HA H5 protein nguyên thể, ví dụ H5 Anhui (RERRRKRGLFGAIAGFIE) (các vị trí axit amin 4-21 của SEQ ID NO: 28), H5 Indo (RESRRKKRGLFGAIAGFIE) (các vị trí axit amin 4-22 của SEQ ID NO: 29), hoặc H5 Vietnam (RERRRKKRGLFGAIAGFIE) (các vị trí axit amin 4-22 của SEQ ID NO: 30), như được thể hiện trên Fig.18C. Đối với H3, trình tự "RNVPEKQTRGIF" (ví dụ các vị trí axit amin 345-356 của SEQ ID NO: 139) có thể được thay thế bằng "GS" (Fig.21E; SEQ ID NO:20), hoặc trình tự H3 "RNVPEKQTR" (ví dụ các vị trí axit amin 345-353 của SEQ ID NO: 139)có thể được thay thế bằng "GSSGSSGSS" (Fig.21F; SEQ ID NO: 21). Đối với chủng B, trình tự "AKLLKERGFFGAIAGFLE" (ví dụ các vị trí axit amin 365-382 của SEQ ID NO: 90) có thể được loại bớt và/hoặc được thay thế bằng trình tự "GG", như được thể hiện trên Fig.21B (SEQ ID NO:17), trình tự "AKLLKERGFFGAIAGFLEG" (ví dụ các vị trí axit amin 365-383 của SEQ ID NO: 90) có thể được thay thế bằng "GG" (như được chỉ ra trong SEQ ID NO: 41), hoặc trình tự "PPAKLLKER" (ví dụ các vị trí axit amin 363-371 của SEQ ID NO: 90) được thay thế bằng "GSSSGSSSG"(Fig.21C; SEQ ID NO: 18).

Các đột biến axit amin có thể được sinh ra bằng cách sử dụng các phương pháp di truyền hoặc hoá học đã được biết đến trong lĩnh vực này. Các phương pháp di truyền có thể bao gồm sự phát sinh đột biến định hướng vị trí, PCR, tổng hợp gen và tương tự. Được dự định rằng các phương pháp làm thay đổi nhóm chuỗi phụ của axit amin bằng các phương pháp khác với việc thao tác di truyền, như sự cải biến hoá học, cũng có thể là hữu ích.

Do đó, trình tự hemagglutinin (HA) của sáng chế có thể bao gồm trình tự vòng phân giải protein được cải biến hoặc vị trí tách, do đó có khả năng phân cắt giảm hoặc không có khả năng phân cắt vòng phân giải protein hoặc vị trí phân cắt bởi proteaza. Các trình tự hemagglutinin polypeptit có thể bao gồm vòng phân giải protein được cải biến hoặc trình tự vị trí tách được cải biến, ví dụ, được nêu trên các hình vẽ Fig.5D, Fig.7D, Fig.8A, Fig.18C, Fig.19, Fig.21B, Fig.21C, Fig.21E, Fig.21F, Fig.24D, Fig.25D, và Fig.26D. Các vị trí tách của trình tự hemagglutinin polypeptit bất kỳ của chủng cúm bất kỳ

có thể được xác định hoặc được dự đoán bằng cách sử dụng bất kỳ trong số các phương pháp được biết đến trong lĩnh vực này, bao gồm việc căn thẳng hàng trình tự (xem ví dụ Fig.15).

Phép phân tích trình tự từ H1, H3 và B HA cho biết H1 có một vị trí phân giải protein monobazơ (Clara typ monobazơ: Q/EXR) mà đứng ngay trước peptit dung hợp (ví dụ các vị trí axit amin 6-8 của SEQ ID NO: 47), trong khi H3 và B HA có 2 vị trí phân giải protein, một vị trí mà được nhận dạng bởi proteaza tương tự Clara (như được tìm thấy trong H1), và vị trí khác được nhận biết bởi trypsin và proteaza tương tự chymotrypsin (P-E/A-K) (ví dụ các vị trí axit amin 3-5 của SEQ ID NO: 48). Trình tự liên ứng đối với việc tách HA được thể hiện trong bảng 1.

Bảng 1: Trình tự liên ứng của vị trí phân giải protein đối với việc tách tiền chất HA_0. Trình tự được nhận dạng bởi Clara tryptaza hoặc trypsin/chimotrypsin có nét in nghiêng và nét đậm một cách tương ứng. Vài chủng HA bao gồm vị trí tách typ Furin polybazơ (RKKR; nội dung rõ rệt, gạch dưới).

H1	NIPSI*QSR*↓GLF	SEQ ID NO: 47
H3	NV**PEK***QTR*↓GIF	SEQ ID NO: 48
H5	TGLRNSPQRESRRKKR↓GLF	SEQ ID NO: 60
B	**PAK**LL*KER*↓GFF	SEQ ID NO: 59

Để tránh được việc tách bằng phân giải protein tiềm ẩn của tiền chất HA0 của HA, chỉ có một vị trí phân giải protein có thể cần phải được cải biến từ trình tự của H1, trong khi trong trường hợp H3 và B, hai vị trí monobazơ có thể cần phải được cải biến.

Ví dụ, vị trí tách thứ nhất của HA_0 của B/Florida và B/Brisbane có thể ví dụ được loại trừ bằng cách thay thế Lys 341 (protein chín đánh số) theo Ile (xem Bảng 2). Vị trí monobazơ thứ hai có thể được huỷ bỏ bằng cách thay thế ba axit amin trước peptit dung hợp, KER (344-346) (ví dụ các vị trí axit amin 17-19 của SEQ ID NO: 36), với NIQ (ví dụ các vị trí axit amin 6-8 của SEQ ID NO: 65). Trình của vài vòng phân giải protein được cải biến của HA được nêu trong bảng 2.

Bảng 2: Sự minh họa của ví dụ về đột biến để phá huỷ việc tách tiền chất HA_0. Vị trí monobazơ có nét in nghiêng (nhận biết tương tự Clara) và nét in đậm (không gạch dưới; tương tự trypsin/chymotrypsin). Đột biến được thể hiện như nét đậm và gạch dưới. Mũi tên thể hiện vị trí tách đối với sự chuyển hoá HA_0 thành HA1-HA2.

Chủng	Trình tự tự nhiên	Huỷ bỏ vị trí tách tiền chất
H5/Indo	TGLRNSPQ**RESRRKKR**↓GLF SEQ ID NO:60	TGLRNSPQ***TETR***GLF SEQ ID NO:61 TGLRNSPQ***TETQ***GLF SEQ ID NO:62
H1/Brisbane	NIPSI*QSR*↓GLF SEQ ID NO:47	NIPSIQS***Q***GLF SEQ ID NO:63
H3/Brisbane	NV**PEK***QTR*↓GIF SEQ ID NO:48	NVPE***I***QT***Q***GIF SEQ ID NO:64
B/Florida, B/Brisbane	**PAK**LL*KER*↓GFF SEQ ID NO:59	PA***I***LL***NIQ***GFF SEQ ID NO:65

Trong ví dụ khác, trình tự bao gồm vòng phân giải protein trong HA0 có thể được thay thế hoặc được loại bớt. Ví dụ, biến thể H3 chứa sự xóa bỏ trình tự RNVPEKQT (ví dụ các vị trí axit amin 345-352 của SEQ ID NO: 122) ở đầu cuối C của HA1 ngoài việc xóa bỏ axit amin đầu cuối N GIFGIA của HA2 (ví dụ các vị trí axit amin 353-356 của SEQ ID NO: 122) được nêu trên Fig.21 E. HA1-HA2 được rút ngắn có thể được liên kết cùng với nhau bởi cầu liên kết GS (các vị trí axit amin 321-322 của SEQ ID NO: 20).

Trong ví dụ khác, vòng chứa vị trí tách phân giải protein trong, ví dụ H3, có thể đã được thay thế bằng cầu liên kết linh hoạt và phần HA2 có thể được để cho nguyên vẹn. Cầu liên kết (GSS)3 có thể được thiết kế để trợ giúp HA1 rút ngắn thành HA2. (xem Fig.21F; SEQ ID NO: 21).

Trong ví dụ khác, HA từ cúm B có thể chứa sự xóa bỏ trình tự AKLLKER (các vị trí axit amin 365-371 của SEQ ID NO: 90) ở đầu cuối C của HA1 ngoài việc xóa bỏ axit amin đầu cuối N GFFGAIAGFLEG của HA2 (các vị trí axit amin 372-383 của SEQ ID NO: 90). HA1-HA2 được rút ngắn có thể được liên kết cùng với nhau bởi cầu liên kết GG (xem ví dụ Fig.21B; Cấu trúc 1008; SEQ ID NO: 17). Sự biểu hiện của cấu trúc này là được thể hiện trên các Fig.13A và Fig.B.

Trong ví dụ khác, HA từ cúm B, vòng chứa vị trí phân giải protein có thể đã được thay thế bằng cầu liên kết linh hoạt và phần HA2 được để cho nguyên vẹn. Cầu liên kết GSSS dài hơn có thể được thiết kế để trợ giúp cho HA1 rút ngắn thành HA2. (xem ví dụ Fig.21C; SEQ ID NO: 18).

Như được thể hiện trên các hình vẽ Fig.13A và Fig.14, HA từ B/Brisbane/60/2008 được biểu hiện kém ở lá *Nicotiana benthamiana* thấm nhiễm nông nghiệp (xem dãy 1008). Tuy nhiên, sự biểu hiện của HA typ B mà đã được cải biến để xóa bỏxóa bỏ vòng phân giải protein (xem dãy 1059, Fig.13A, Fig.14) dẫn đến sự biểu hiện gia tăng. Hơn nữa, sự đồng biểu hiện của HA-typ B với M2 từ A/New Caledonia/20/99, dẫn đến sự gia tăng biểu hiện HA (xem các dãy "1008+1261"; và 1059+1261"). Sự đồng biểu hiện của HA typ B bao gồm việc xóa bỏ trong vòng phân giải protein, với M2 từ A/Puerto Rico/8/34 cũng dẫn đến sự biểu hiện gia tăng (1059+859; Fig.14).

Theo cách tương tự, việc xóa bỏ vòng phân giải protein trong H5/Indo, và sự thay thế bằng "GG" (cấu trúc 928; xem Fig.25D; SEQ ID NO: 85), "TETR" (cấu trúc 676; cũng xem Fig.19, Fig.24D) hoặc "TETQ" (cấu trúc 766; cũng xem các Fig.19, Fig.25D) trình tự dẫn đến mức biểu hiện mà khớp hoặc gia tăng so với mức biểu hiện được quan sát với H5/Indo nguyên thể (cấu trúc 489; xem Fig.20 và Fig.23).

Như được thể hiện trên Fig.13B, bằng cách xóa bỏ vòng phân giải protein của HA0 (trình tự được thể hiện trên Fig.21B; SEQ ID NO: 17), HA0 protein thu được thể hiện hoạt tính gia tăng như được thể hiện bởi khả năng hemagglutin hoá lớn hơn, khi so với protein HA mà không có vòng phân giải protein của nó được loại bỏ.

Bằng việc gia tăng hoạt tính, có nghĩa là gia tăng khả năng hemagglutin hoá khoảng từ 2% đến 100%, hoặc lượng bất kỳ giữa chúng như được xác định bằng cách sử dụng các kỹ thuật tiêu chuẩn trong lĩnh vực này, ví dụ, từ 10% đến 50% hoặc trị số bất kỳ giữa chúng ví dụ khoảng 2, 5, 8, 10, 12, 15, 18, 20, 22, 24, 25, 26, 28, 30, 32, 34, 35, 36, 38, 40, 42, 44, 45, 46, 48, 50, 52, 54, 55, 56, 58, 60, 65, 70, 75, 80, 85, 90, 95, hoặc 100%, hoặc lượng bất kỳ giữa chúng, khi so với hoạt tính của cùng protein HA mà không có vòng phân giải protein của nó được loại bỏ.

Sáng chế cũng bao gồm trình tự nucleotit mã hoá HA được cải biến từ ví dụ H1,H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16 hoặc typ B HA được cải biến hoặc trình tự nucleotit bất kỳ mà lai với H1,H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16 hoặc typ B HA trong các điều kiện nghiêm ngặt hoặc trình tự nucleotit mà lai trong các điều kiện lai nghiêm ngặt với sự hỗ trợ của H1,H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16 hoặc typ B HA, trong đó trình tự nucleotit mã hoá hemagglutinin protein mà khi được biểu hiện tạo thành VLP, và VLP gây ra việc sản xuất kháng thể. Ví dụ, sự biểu hiện của trình tự nucleotit với tế bào thực vật tạo thành VLP, và VLP có thể được sử dụng để tạo ra kháng thể mà có khả năng gắn kết HA, bao gồm HA chín từ B hoặc H3. VLP, khi được dùng cho đối tượng, gây ra đáp ứng miễn dịch. Tốt hơn là, VLP gây ra việc sản xuất kháng thể và VLP, khi được dùng cho đối tượng, gây ra đáp ứng miễn dịch.

Ví dụ, sự biểu hiện của trình tự nucleotit với tế bào thực vật tạo thành VLP, và VLP có thể được sử dụng để tạo ra kháng thể mà có khả năng gắn kết protein virut như ví dụ HA, bao gồm nhưng không chỉ giới hạn ở HA0, HA0 protein với vòng phân giải protein của nó được loại bớt hoặc được cải biến, HA1 hoặc HA2 của một hoặc nhiều typ cúm hoặc typ phụ, ví dụ nhưng không chỉ giới hạn ở typ phụ H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, typ B HA. VLP, khi được dùng cho đối tượng, gây ra đáp ứng miễn dịch.

Việc lai trong các điều kiện lai nghiêm ngặt là đã được biết đến trong lĩnh vực này (xem ví dụ ấn phẩm: Current Protocols in Molecular Biology, Ausubel et al., eds. 1995 và phần bổ sung; Maniatis *et al.*, in Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory, 1982; Sambrook và Russell, in Molecular Cloning: A Laboratory Manual, 3rd edition 2001). Ví dụ về một trong số các điều kiện lai nghiêm ngặt này có thể là lai khoảng từ 16 đến 20 giờ trong 4 X SSC ở 65°C, tiếp theo rửa trong 0,1 X SSC ở 65°C trong một giờ hoặc 2 rửa trong 0,1 X SSC ở 65°C, mỗi lần trong 20 hoặc 30 phút. Theo cách khác, điều kiện lai nghiêm ngặt lấy làm ví dụ có thể là qua đêm (từ 16 đến 20 giờ) trong 50% formamit, 4 X SSC ở 42°C, tiếp theo rửa trong 0,1 X SSC ở 65°C trong một giờ hoặc 2 lần trong 0,1 X SSC ở 65°C, mỗi lần trong 20 hoặc 30 phút hoặc qua đêm

(từ 16 đến 20 giờ) hoặc lai trong chất đệm phosphat trong nước Church (7% SDS; 0,5M chất đệm $NaPO_4$ độ pH=7,2; 10mM EDTA) ở 65°C, với 2 lần rửa ở 50°C trong 0,1 X SSC, 0,1% SDS trong 20 hoặc 30 phút mỗi lần hoặc 2 lần rửa ở 65°C trong 2 X SSC, 0,1% SDS trong 20 hoặc 30 phút mỗi lần.

Ngoài ra, sáng chế bao gồm trình tự nucleotit mà được khác biệt do có khoảng 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% hoặc lượng bất kỳ giữa chúng, giống nhau về trình tự hoặc tương tự về trình tự, với trình tự nucleotit mã hoá H1,H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16 hoặc typ B HA, trong đó trình tự nucleotit mã hoá hemagglutinin protein (HA được cải biến) với trình tự vòng phân giải protein được cải biến hoặc vị trí tách mà làm giảm hoặc huỷ bỏ việc tách vòng phân giải protein hoặc vị trí phân cắt bởi proteaza. Khi trình tự nucleotit mã hoá HA được cải biến được biểu, nó tạo thành VLP, và VLP gây ra việc sản xuất kháng thể. Ví dụ, sự biểu hiện của trình tự nucleotit với tế bào thực vật tạo thành VLP, và VLP có thể được sử dụng để tạo ra kháng thể mà có khả năng gắn kết HA, bao gồm HA (HA0) không được xử lý hoặc không được xử lý trong đó vòng phân giải protein đã được loại bỏ. VLP, khi được dùng cho đối tượng, gây ra đáp ứng miễn dịch.

Ngoài ra, sáng chế bao gồm các trình tự nucleotit mà khác biệt do có khoảng 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% hoặc lượng bất kỳ giữa chúng, giống nhau về trình tự hoặc tương tự về trình tự, với trình tự nucleotit của SEQ ID NO: 43, 91, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 137, 140, 144, 151, 158, 165, trong đó trình tự nucleotit mã hoá protein HA được cải biến mà khi được biểu hiện tạo thành VLP, và mà VLP gây ra việc sản xuất kháng thể mà có khả năng gắn kết HA, bao gồm HA (HA0) không được xử lý hoặc không được xử lý trong đó vòng phân giải protein đã được loại bớt hoặc được cải biến. VLP, khi được dùng cho đối tượng, gây ra đáp ứng miễn dịch.

Hơn nữa, sáng chế bao gồm trình tự axit amin mà khác biệt do có khoảng 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% hoặc lượng bất kỳ giữa chúng, sự giống nhau về trình tự hoặc tương tự về trình tự, với trình tự axit amin của SEQ ID NO: 17, 18, 20, 21, 41, 58, 77, 81, 85, 92, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 134, 143, 147, 154, 161,168, 194 và 199, trong đó trình tự axit amin mã hoá protein HA được

cải biến mà khi được biểu hiện tạo thành VLP, và mà VLP gây ra việc sản xuất kháng thể mà có khả năng gắn kết HA, bao gồm HA (HA0) không được xử lý hoặc không được xử lý trong đó vòng phân giải protein đã được loại bớt hoặc được cải biến. VLP, khi được dùng cho đối tượng, gây ra đáp ứng miễn dịch.

Sự giống nhau về trình tự hoặc sự tương tự về trình tự có thể được xác định bằng cách sử dụng chương trình so sánh trình tự nucleotit, như trình tự được nêu trong DNASIS (ví dụ, bằng cách sử dụng, nhưng không chỉ giới hạn ở, các tham số sau đây: điểm phạt GAP 5, #của đường chéo đỉnh 5, điểm phạt GAP 10 cố định, k-tuple 2, khoảng trống nổi 10 và kích cỡ cửa sổ 5). Tuy nhiên, các phương pháp khác về việc căn thẳng hàng trình tự để so sánh là đã được biết đến trong lĩnh vực này ví dụ thuật toán của Smith & Waterman (1981, Adv. Appl. Math. 2:482), Needleman & Wunsch (J. Mol. Biol. 48:443, 1970), Pearson & Lipman (1988, Proc. Nat'l. Acad. Sci. USA 85:2444), và bằng việc thực hiện máy tính hoá các thuật toán này (ví dụ GAP, BESTFIT, FASTA, và BLAST), hoặc bằng việc căn thẳng hàng bằng tay và sự thẩm định bằng mắt thường. Ví dụ về sự căn thẳng hàng trình tự của HA từ các chủng cúm khác nhau có thể được tìm thấy trên Fig.24.

Ví dụ, nhưng không chỉ giới hạn ở, trình tự nucleotit mã hoá:

- typ B HA với vòng phân giải protein được cải biến như được xác định bởi SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 58, SEQ ID NO: 96, SEQ ID NO: 98, SEQ ID NO: 100, SEQ ID NO: 102, SEQ ID NO: 104, SEQ ID NO: 106, SEQ ID NO: 108, SEQ ID NO: 110, SEQ ID NO: 112, SEQ ID NO: 114 và SEQ ID NO: 168, hoặc trình tự nucleotit mã hoá typ B HA bao gồm vùng vòng phân giải protein được cải biến như được xác định trong SEQ ID NO: 65, SEQ ID NO: 72, SEQ ID NO:73, SEQ ID NO:95, SEQ ID NO: 97, SEQ ID NO: 99, SEQ ID NO: 101, SEQ ID NO: 103, SEQ ID NO: 105, SEQ ID NO: 107, SEQ ID NO: 109, SEQ ID NO: 111, SEQ ID NO: 113 và SE ID NO: 165.

- H1 với vòng phân giải protein được cải biến bao gồm các trình tự bao gồm vị trí tách được cải biến như được xác định bởi SEQ ID NO: 63.

- H2 với vòng phân giải protein được cải biến bao gồm các trình tự bao gồm vị trí tách được cải biến như được xác định bởi SEQ ID NO: 134.

- H3 với vòng phân giải protein được cải biến bao gồm các trình tự được xác định bởi SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO 143, SEQ ID NO: 147 hoặc bao gồm vị trí tách được cải biến như được xác định bởi SEQ ID NO: 64.

- H5 với vòng phân giải protein được xóa bỏ bao gồm các trình tự bao gồm vị trí tách được cải biến như được xác định bởi SEQ ID NO: 61, SEQ ID NO:62, SEQ ID NO: 69, SEQ ID NO: 70, SEQ ID NO: 71.

- H7 với vòng phân giải protein được cải biến bao gồm các trình tự bao gồm vị trí tách được cải biến như được xác định bởi SEQ ID NO: 154 hoặc trình tự nucleotit mã hoá typ H7 HA bao gồm vùng vòng phân giải protein được cải biến như được xác định trong SEQ ID NO: 151.

- H9 với vòng phân giải protein được cải biến bao gồm các trình tự bao gồm vị trí tách được cải biến như được xác định bởi SEQ ID NO: 161 hoặc trình tự nucleotit mã hoá typ H9 HA bao gồm vùng vòng phân giải protein được cải biến như được xác định trong SEQ ID NO: 158.

Sáng chế đề cập đến việc sử dụng protein HA bao gồm miền xuyên màng và bao gồm các miền HA1 và HA2, ví dụ protein HA có thể là HA0 hoặc HA được xử lý bao gồm HA1 và HA2. protein HA có thể được sử dụng trong việc tạo ra hoặc hình thành VLP bằng cách sử dụng hệ biểu hiện thực vật hoặc tế bào thực vật.

Yếu tố khuếch đại và Yếu tố tăng cường/Yếu tố điều hoà

Trong ví dụ khác, protein HA được cải biến có thể được biểu hiện trong hệ biểu hiện mà bao gồm yếu tố khuếch đại và/hoặc yếu tố hoặc vùng điều hoà (cũng được đề cập trong bản mô tả này là yếu tố tăng cường). Ví dụ, yếu tố khuếch đại từ geminivirut như ví dụ, yếu tố khuếch đại từ virut vàng lùn đậu (BeYDV) có thể được sử dụng để biểu hiện HA được cải biến. BeYDV thuộc về chi Mastrevirut được làm thích ứng với thực vật có hai lá mầm. BeYDV là monopartit có hệ gen ADN mạch vòng sợi đơn và có thể sao chép

thành số lượng bản sao rất lớn bằng cơ chế lăn vòng. Các hệ vật truyền replicon ADN thu được từ BeYDV đã được sử dụng để sản xuất protein sản lượng cao nhanh chóng ở thực vật.

Như được sử dụng trong bản mô tả này, cụm từ "yếu tố khuếch đại" đề cập đến đoạn axit nucleic bao gồm ít nhất một phần của một hoặc nhiều vùng liên gen dài (LIR) của hệ gen geminivirut. Như được sử dụng trong bản mô tả này, "vùng liên gen dài" đề cập đến vùng của vùng liên gen dài mà chứa vị trí gắn kết sao chép có khả năng làm trung gian cắt và sao chép bởi geminivirut Rep protein. Theo một số khía cạnh, đoạn axit nucleic bao gồm một hoặc nhiều LIR, còn có thể bao gồm vùng liên gen ngắn (SIR) của hệ gen geminivirut. Như được sử dụng trong bản mô tả này, "vùng liên gen ngắn" đề cập đến sợi bổ trợ (IR ngắn (SIR) của Mastrevirut). Yếu tố khuếch đại thu được từ geminivirut thích hợp bất kỳ có thể được sử dụng trong bản mô tả này. Xem, ví dụ công bố đơn quốc tế số WO2000/20557; WO2010/025285; ấn phẩm: Zhang X. et al. (2005, *Biotechnology and Bioengineering*, Vol. 93, 271-279), Huang Z. et al. (2009, *Biotechnology and Bioengineering,* Vol. 103, 706-714), Huang Z. et al.(2009, *Biotechnology and Bioengineering,* Vol. 106, 9-17)). Nếu nhiều hơn một LIR được sử dụng trong cấu trúc, ví dụ hai LIR, thì đoạn khởi đầu, vùng CMPV-HT và trình tự axit nucleic quan tâm và đoạn kết thúc được đặt trong dấu ngoặc bởi mỗi trong số hai LIR.

Như được mô tả trong bản mô tả này, vật truyền thu được từ đồng phân phối của virut vàng lùn đậu (BeYDV) và vật truyền cung cấp Rep/RepA, bằng việc thấm lọc nông nghiệp của lá *Nicotiana benthamiana* dẫn đến sự khuếch đại replicon hữu hiệu và sản xuất protein mạnh. Phép phân tích thấm tách Western của phần chiết protein từ thực vật được biến nạp với cấu trúc gen điều khiển sự biểu hiện của cúm B HA được cải biến (từ B/Brisbane/60/2008) có hoặc không có vòng phân giải protein được loại bỏ (xem Fig.17A đối với các cấu trúc) và với sự có mặt hoặc không có mặt của yếu tố khuếch đại BeYDV (cấu trúc số 1059 và 1039) thể hiện rằng với sự không có mặt của BeYDV, không có sự kết tụ cúm B HA có thể được phát hiện (Fig.17B), khi yếu tố điều hoà là CPMV-HT.

Như được thể hiện trên Fig.17B, sự biểu hiện của HA từ B/Brisbane/60/2008 với vòng phân giải protein được loại bỏ với sự không có mặt của BeYDV không dẫn đến sự

biểu hiện có thể phát hiện bởi phép phân tích thấm tách Western (xem dãy 1039 trên Fig.17B). Tuy nhiên, sự biểu hiện của HA typ B với vòng phân giải protein được loại bỏ với sự có mặt của yếu tố khuếch đại BeYDV, dẫn đến sự biểu hiện gia tăng (xem dãy 1059). Tương tự, với sự không có mặt của BeYDV, sự đồng biểu hiện của thể đột biến HA-typ B bao gồm việc xóa bỏ trong vòng phân giải protein, với M2 từ A/New Caledonia/20/99, không dẫn đến sự biểu hiện HA có thể phát hiện (xem các dãy "1039+1261" trên Fig.17B). Sự đồng biểu hiện của thể đột biến HA typ B bao gồm việc xóa bỏ trong vòng phân giải protein với sự có mặt của BeYDV, với M2 từ A/New Caledonia/20/99 mặt khác dẫn đến sự biểu hiện gia tăng (xem dãy "1059+1261"; Fig.17B).

Tuy nhiên, sự có mặt của BeyDV là không được đòi hỏi khi yếu tố tăng cường là có mặt trong hệ biểu hiện và khi yếu tố tăng cường không phải là CPMV-HT. Như ví dụ được thể hiện trên Fig.29A, sự biểu hiện của các chủng B HA khác nhau dưới sự kiểm soát của yếu tố tăng cường, ví dụ CPMV 160, CPMV160+ hoặc CPMV HT+, dẫn đến việc sản xuất protein HA mà thể hiện độ chuẩn hemagglutin hoá (HMG) gia tăng với sự không có mặt của BeYDV.

Do đó, protein HA đột biến (được cải biến) có thể được biểu hiện với sự không có mặt của yếu tố khuếch đại, như yếu tố khuếch đại được tạo thành chủ yếu từ geminivirut ví dụ BeYDV, nhưng với sự có mặt của yếu tố tăng cường, ví dụ CPMV 160, CPMV160+ hoặc CPMV HT+.

Thể đột biến (HA được cải biến) có thể được biểu hiện với sự có mặt của yếu tố tăng cường, ví dụ CPMV 160, CPMV160+ hoặc CPMV HT+, nhưng với sự.không có mặt hoặc với sự có mặt của yếu tố khuếch đại, ví dụ BeYDV. Như được thể hiện trên Fig.28B, Fig.28C và Fig.28F HA đột biến (được cải biến) có thể được biểu hiện với sự có mặt của yếu tố tăng cường, có hoặc không có mặt của yếu tố khuếch đại. Do đó, sáng chế cũng được dẫn đến sự biểu hiện của HA đột biến (được cải biến) với sự có mặt của yếu tố tăng cường và tuỳ ý yếu tố khuếch đại.

Cấu trúc HA bao gồm yếu tố tăng cường (CMPV HT+ hoặc CMPV 160+) và vòng phân giải protein được thay thế bằng cầu liên kết GG (vòng phân giải protein được xóa bỏ) thể hiện sự biểu hiện gia tăng khi so với các cấu trúc kiểu hoang dại hoặc HA bao gồm CPMV HT (Fig.28A, H3 Per; Fig.28B, B Malaysia; Fig.28C, H9 HK; Fig.29D, B Mass; Fig.28E, H2 Sin).

Fig.29A thể hiện bản tổng kết dữ liệu đối với độ chuẩn hemagglutin hoá của protein HA được cải biến được tạo ra ở thực vật bao gồm yếu tố tăng cường được tạo thành chủ yếu từ CPMV HT, CPMV HT+, CPMV 160 hoặc CPMV160+,.được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá HA được cải biến với vòng phân giải protein được xóa bỏ (cầu liên kết GG) hoặc HA nguyên thể. Trong hầu hết các trường hợp, sự biểu hiện (được xác định như độ chuẩn hemagglutin hoá) là cao hơn đối với cấu trúc được tạo thành chủ yếu từ CPMV HT+, CPMV 160 hoặc CPMV160+ cho biết mức biểu hiện đáng kể.

Yếu tố tăng cường có thể được sử dụng để đạt được mức biểu hiện tạm thời cao của protein HA đột biến (được cải biến) với vòng phân giải protein được cải biến. Yếu tố tăng cường có thể được tạo thành chủ yếu từ ARN virut thực vật, bao gồm comovirut, như *Cowpea mosaic virus* (CPMV; xem, ví dụ, WO2007/135480; WO2009/087391; US 2010/0287670, Sainsbury F. et al., 2008, *Plant Physiology;* 148: 121-1218; Sainsbury F. et al., 2008, *Plant Biotechnology Journal;* 6: 82-92; Sainsbury F. et al., 2009, *Plant Biotechnology Journal;* 7: 682-693; Sainsbury F. et al. 2009, *Methods in Molecular Biology, Recombinant Proteins From Plants*, vol. 483: 25-39).

*CPMV 160 (*CPMVX) *và CPMV 160+* (CPMVX+)

Theo một phương án, yếu tố tăng cường là "CPMVX" (cũng được đề cập đến là "CPMV 160") và/hoặc "CPMVX+" (cũng được đề cập đến là "CPMV 160+") như được mô tả trong US 61/925,852.

Vùng tăng cường sự biểu hiện "CPMVX" bao gồm vùng không dịch mã 5' (UTR) virut khảm đậu đũa comovirut (CPMV). 5'UTR từ nucleotit từ 1 đến 160 của trình tự CPMV RNA-2 (SEQ ID NO: 93), bắt đầu ở vị trí khởi đầu phiên mã đến codon bắt đầu

khởi đầu trong khung thứ nhất (ở vị trí 161), mà có tác dụng như vị trí khởi đầu để tạo ra hơn nữa hai protein đồng kết thúc carboxy được mã hoá bởi đoạn hệ gen comovirut kiểu hoang dại. Hơn nữa, vị trí khởi đầu 'thứ ba' ở (hoặc tương ứng với) vị trí 115 trong trình tự hệ gen CPMV RNA-2 cũng có thể được đột biến, loại bớt hoặc theo cách khác được thay thế. Nó đã được thể hiện rằng việc loại bỏ AUG 115 ngoài việc loại bỏ AUG 161 làm tăng sự biểu hiện khi tổ hợp với M protein không hoàn thiện (Sainsbury and Lomonossoff, 2008, *Plant Physiology;* 148: 1212-1218; WO 2009/087391).

CPMVX bao gồm X nucleotit của SEQ ID NO:93, trong đó X=160, 155, 150, hoặc 114 của SEQ ID NO:93, hoặc trình tự mà bao gồm khoảng từ 80% đến 100% trình tự tương tự với CPMVX, trong đó X=160, 155, 150, hoặc 114 của SEQ ID NO:93. Vùng tăng cường sự biểu hiện này thường được đề cập đến CPMVX (xem Fig.26A).

Vùng tăng cường sự biểu hiện CPMVX, trong đó X=160, gồm có nucleotit từ 1 đến 160 của SEQ ID NO: 93:

```
1   tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa
ccaaaccttc

61 ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc
ttgcgtgagc

121 gatcttcaac gttgtcagat cgtgcttcgg caccagtaca (SEQ ID
NO:93)
```

Trình tự tăng cường CPMVX còn có thể được dung hợp với trình tự nhồi, trong đó CMPVX bao gồm X nucleotit của SEQ ID NO: 93, trong đó X=160, 155, 150, hoặc 114 của SEQ ID NO: 93, hoặc trình tự mà bao gồm khoảng từ 80 đế 100 % trình tự tương tự với CPMVX, trong đó X=160, 155, 150, hoặc 114 của SEQ ID NO:93, và trình tự nhồi bao gồm từ 1 đến 100 nucleotit được dung hợp với đầu 3' của trình tự CMPVX. Ví dụ, trình tự nhồi có thể bao gồm khoảng 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, hoặc 100 nucleotit, hoặc số lượng bất kỳ của các nucleotit nằm trong đó.

Nếu trình tự CMPVX bao gồm mảnh nhồi, thì vùng tăng cường sự biểu hiện này có thể được đề cập đến như CPMVX+ (xem Fig.26A), trong đó X=160, 155, 150, 114 của

SEQ ID NO: 93, nó cũng được đề cập đến như CMPVX bao gồm trình tự nhồi hoặc nó có thể được đề cập đến như CPMV160+; CPMV155+; CPMV150+; CPMV114+, khi X-160, 155, 150, hoặc 114, một cách tương ứng. Cấu trúc bao gồm CPMVX mà không bao gồm trình tự nhồi có thể được gọi là CPMVX+, trong đó X=160, 155, 150, 114 của SEQ ID NO:93, và trong đó trình tự nhồi có chiều dài 0 nucleotit.

Trình tự nhồi có thể được cải biến bằng việc cắt cụt, xóa bỏ hoặc thay thế trình tự CMPV5'UTR nguyên thể mà được định vị 3' vào nucleotit 160. Trình tự nhồi được cải biến có thể được xóa bỏ, thay thế, cắt cụt hoặc rút ngắn khi so với trình tự nhồi ban đầu hoặc không được cải biến (nghĩa là nguyên thể) kết hợp với 5'UTR (như được mô tả trong ấn phẩm: Sainsbury F., and Lomonossoff G.P., 2008, Plant Physiol. 148: pp. 1212-1218). Trình tự nhồi có thể bao gồm một hoặc nhiều vị trí giới hạn (đa liên kết, vị trí tạo đa dòng, một hoặc nhiều vị trí tạo dòng), một hoặc nhiều trình tự kozak thực vật, một hoặc nhiều trình tự liên kết, một hoặc nhiều vị trí tái tổ hợp hoặc tổ hợp của chúng. Ví dụ, mà không được xét đến là giới hạn, trình tự nhồi có thể bao gồm trong dãy, vị trí tạo đa dòng có chiều dài mong muốn được dung hợp với trình tự kozak của thực vật. Trình tự nhồi không bao gồm trình tự nucleotit từ trình tự 5'UTR nguyên thể mà được định vị 3' vào nucleotit 160 của CPMV 5'UTR nguyên thể, ví dụ nucleotit 161 đến 512 như được thể hiện trên Fig.1 của ấn phẩm: Sainsbury F., and Lomonossoff G.P. (2008, Plant Physiol. 148: pp. 1212-1218) hoặc nucleotit 161 đến 509 của SEQ ID NO:4. Tức là, M protein không hoàn thiện có mặt trong trình tự CPMV HT của kỹ thuật hiện nay (Fig.1; của ấn phẩm: Sainsbury F., and Lomonossoff G.P., 2008) được loại bỏ ra khỏi 5'UTR trong sáng chế.

Các trình tự liên ứng Kozak của thực vật được biết trong lĩnh vực này (xem ví dụ ấn phẩm: Rangan et al. Mol. Biotechnol., 2008, July 39(3), pp. 207-213). Các trình tự Kozak xuất hiện tự nhiên và tổng hợp có thể được sử dụng trong vùng tăng cường sự biểu hiện hoặc có thể được dung hợp với trình tự nucleotit quan tâm như được mô tả trong bản mô tả này.

Trình tự kozak thực vật có thể là trình tự kozak thực vật đã được biết đến bất kỳ (xem ví dụ ấn phẩm: L. Rangan et. al. Mol. Biotechnol.2008), bao gồm nhưng không chỉ giới hạn ở các trình tự liên ứng của thực vật sau đây:

caA(A/C)a (SEQ ID NO:174; giới thực vật)

aaA(A/C)a (SEQ ID NO:175; thực vật hai lá mầm)

aa(A/G)(A/C)a (SEQ ID NO:176; arabidopsis)

Trình tự kozak của thực vật cũng được chọn từ nhóm gồm có:

AGAAA (SEQ ID NO: 177)

AGACA (SEQ ID NO: 178)

AGGAA (SEQ ID NO: 179)

AAAAA (SEQ ID NO: 180)

AAACA (SEQ ID NO: 181)

AAGCA (SEQ ID NO: 182)

AAGAA (SEQ ID NO: 183)

AAAGAA (SEQ ID NO: 184)

AAAGAA (SEQ ID NO: 185)

(A/-)A(A/G)(A/G)(A/C)A. (SEQ ID NO: 186; trình tự liên ứng)

Vùng tăng cường sự biểu hiện CPMVX, hoặc CPMVX+, có thể được liên kết theo cách hoạt động được ở đầu 5' của trình tự tăng cường với vùng điều hoà mà có hoạt tính ở thực vật, và được liên kết theo cách hoạt động được với trình tự nucleotit quan tâm ở đầu 3' của vùng tăng cường sự biểu hiện (Fig.26A), để điều khiển sự biểu hiện của trình tự nucleotit quan tâm với vật chủ thực vật.

CPMV HT+ ,CPMV HT+[WT115], CPMV HT+ [511]

Theo phương án khác, tếu tố tăng cường là "CPMV HT+" như được mô tả trong công bố patent Mỹ số US 61/971,274. Vùng tăng cường sự biểu hiện "CPMV HT+" (xem Fig.27A) bao gồm vùng không dịch mã comovirut 5' (UTR) và trình tự nhồi cải biến, kéo dài hoặc cắt cụt.

Hệ biểu hiện thực vật bao gồm trình tự axit nucleic thứ nhất bao gồm vùng điều hoà, được liên kết theo cách hoạt động được với một hoặc nhiều hơn một vùng tăng cường sự biểu hiện như được mô tả trong bản mô tả này (ví dụ CPMV HT+, CPMV HT+[WT115], CPMV HT+ [511]), và trình tự nucleotit mã hoá HA được cải biến cũng được đề xuất. Hơn nữa, axit nucleic bao gồm trình tự đoạn khởi đầu (vùng điều hoà), vùng tăng cường sự biểu hiện (ví dụ CPMV HT+ hoặc CPMV HT+[WT115]) bao gồm comovirut 5'UTR và trình tự nhồi với trình tự kozak của thực vật được dung hợp với một hoặc nhiều trình tự axit nucleic mã hoá HA được cải biến là được mô tả. Axit nucleic còn có thể bao gồm trình tự bao gồm vùng không dịch mã comovirut 3' (UTR), ví dụ, plastocyanin 3' UTR, hoặc 3'UTR khác có hoạt tính ở thực vật, và trình tự kết thúc, ví dụ trình tự kết thúc NOS, được liên kết theo cách hoạt động được với đầu 3' của trình tự nucleotit mã hoá HA được cải biến (được đề cập đến như nucleotit quan tâm trên Fig.27A), sao cho trình tự nucleotit mã hoá HA được cải biến xen vào ngược dòng từ vùng không dịch mã 3' comovirut (UTR), plastoxyanin 3' UTR, hoặc trình tự 3'UTR khác.

SEQ ID NO:173 bao gồm "CPMV HT" vùng tăng cường sự biểu hiện như được biết trong lĩnh vực này (ví dụ Fig.1 của ấn phẩm: Sainsbury and Lomonossoff 2008, Plant Physiol. 148: pp. 1212-1218). CPMV HT bao gồm trình tự 5'UTR từ nucleotit từ 1 đến 160 của SEQ ID NO:173 với nucleotit được cải biến ở vị trí 115 (cgt), và M protein không hoàn thiện với nucleotit được cải biến ở vị trí 162 (acg), và thiếu trình tự kozak của thực vật (5'UTR: nucleotit từ 1 đến 160; M protein không hoàn thiện gạch dưới, nucleotit từ 161 đến 509). SEQ ID NO:173 cũng bao gồm vị trí tạo đa dòng (chữ nghiêng, nucleotit từ 510 đến 528) mà không có mặt trong trình tự CPMV HT của kỹ thuật hiện nay:

1 tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc

61 ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc ttgcgtgagc

121 gatcttcaac gttgtcagat cgtgcttcgg caccagtaca *acg*ttttctt tcactgaagc

181 gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg tgtacttgtc

241 ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct gttcagcccc

301 atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct acttctgctt

361 gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt tctataagaa

421 atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga gaaagattgt

481 taagcttctg tatattctgc ccaaatttgt *cgggccc* SEQ ID NO: 173

CPMV HT+ với trình tự liên ứng kozak của thực vật được nêu trong SEQ ID NO:187 (nucleotit từ 1 đến 160, 5'UTR, bao gồm ATG được cải biến ở vị trí 115 (GTG) nét chữ đậm dưới và chữ nghiêng; mảnh nhồi bao gồm: M protein không hoàn thiện gạch dưới, nucleotit từ 161 đến 509, với nucleotit được cải biến ở 162 (ACG); vị trí tạo đa dòng, chữ nghiêng, nucleotit từ 510 đến 528; và trình tự liên ứng kozak của thực vật, chữ cái hoa và đậm, nucleotit từ 529 đến 534).

1 tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc

61 ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc ttgc***gtg***agc

121 gatcttcaac gttgtcagat cgtgcttcgg caccagtaca ***acg***ttttctt tcactgaagc

181 gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg tgtacttgtc

241 ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct gttcagcccc

301 atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct acttctgctt

361 gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt tctataagaa

421 atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga gaaagattgt

481 taagcttctg tatattctgc ccaaatttg*t tcgggcccaa taccgcgg**(A/-)A(A/G)*

(A/G)(A/C)A (SEQ ID NO:187)

SEQ ID NO:188 ("CPMV HT+ 511") bao gồm đoạn trình tự nguyên thể của hệ gen CPMV RNA 2 từ nucleotit 1 đến 154. Trình tự 5'UTR từ nucleotit 1 đến 511 của SEQ ID NO:188 bao gồm các trình tự "atg" được cải biến ở vị trí 115 ("g" thay thế cho "a"; chữ nghiêng đậm) và 162 ("c" thay thế cho "t"; chữ nghiêng đậm), và M protein không hoàn thiện (gạch dưới) từ nucleotit từ 161 đến 511. CPMV HT+ 511 bao gồm trình tự liên ứng M protein kozak của thực vật (nucleotit từ 508 đến 511; đậm):

1 tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc

61 ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc ttgc***gtg***agc

121 gatcttcaac gttgtcagat cgtgcttcgg caccagtaca ***acg***ttttctt tcactgaagc

181 gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg tgtacttgtc

241 ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct gttcagcccc

301 atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct acttctgctt

361 gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt tctataagaa

421 atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga gaaagattgt

481 taagcttctg tatattctgc ccaaatttga a... SEQ ID NO: 188

Ví dụ không hạn chế khác của trình tự tăng cường CPMV HT+ được đưa ra bởi trình tự SEQ ID NO:189 (CPMV HT+[WT115]). Cátxét biểu hiện hoặc vật truyền bao gồm CPMV HT+ và bao gồm vùng điều hoà thực vật trong sự kết hợp thao tác với trình tự vùng tăng cường sự biểu hiện của SEQ ID NO: 189, và vị trí khởi đầu phiên mã (ATG) ở đầu 3' dung hợp với trình tự nucleotit mã hoá HA được cải biến cũng là một phần của sáng chế.

SEQ ID NO: 189 (CPMV HT+[WT115]) nucleotit từ 1 đến 160, 5'UTR, với ATG ở vị trí từ 115 đến 117, nét đậm dưới; đoạn nhồi bao gồm: M protein không hoàn thiện gạch dưới, nucleotit từ 161 đến 509; với ATG được cải biến ở vị trí từ 161 đến 153 nét đậm dưới và gạch dưới, vị trí tạo đa dòng, chữ nghiêng, nucleotit từ 510 đến 528; và trình tự kozak của thực vật, chữ cái hoa nét đậm, nucleotit từ 529 đến 534).

1 tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc

61 ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc ttgc***atg***agc

121 gatcttcaac gttgtcagat cgtgcttcgg caccagtaca ***acg***ttttctt tcactgaagc

181 gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg tgtacttgtc

241 ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct gttcagcccc

301 atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct acttctgctt

361 gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt tctataagaa

421 atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga gaaagattgt

481 taagcttctg tatattctgc ccaaatttgt *tcgggcccaa taccgcgg**AG AAAA***

(SEQ ID NO:189)

Trình tự kozak của thực vật của SEQ ID NO:189 có thể là trình tự kozak bất kỳ của thực vật, bao gồm nhưng không chỉ giới hạn, ở một trong số các trình tự SEQ ID NO: từ 174 đến 186.

"Protein khảm"

HA được cải biến còn có thể là protein khảm. Thuật ngữ "protein virut khảm" hoặc "polypeptit virut khảm", cũng được đề cập đến là "protein khảm" hoặc "polypeptit khảm", hoặc "HA khảm" có nghĩa là protein hoặc polypeptit mà bao gồm trình tự axit amin từ hai hoặc nhiều hơn hai nguồn, ví dụ nhưng không chỉ giới hạn ở, hai hoặc nhiều typ cúm hoặc typ phụ hoặc cúm có nguồn gốc khác nhau, mà được dung hợp như polypeptit đơn lẻ. Protein hoặc polypeptit khảm có thể bao gồm peptit tín hiệu mà là giống với hoặc khác loại với, phần còn lại của polypeptit hoặc protein. Protein khảm hoặc polypeptit khảm có thể được tạo ra như bản sao từ trình tự nucleotit khảm, và tổng hợp tiếp theo và nếu cần, có thể kết hợp để tạo thành protein multime. Do đó, protein khảm hoặc polpypeptit khảm cũng bao gồm protein hoặc polypeptit bao gồm cấu trúc dưới đơn vị mà được kết hợp qua các cầu disulfua (nghĩa là protein multime). Ví dụ, polypeptit khảm bao gồm trình tự axit amin từ hai hoặc nhiều hơn hai nguồn có thể được xử lý thành các dưới đơn vị và các dưới đơn vị được kết hợp qua cầu disulfua để tạo ta protein khảm hoặc polypeptit khảm. protein HA khảm cũng có thể bao gồm protein kháng nguyên hoặc mảnh của nó của virut cúm thứ nhất và phức hệ miền xuyên màng (TDC) từ HA virut cúm thứ hai, bao gồm miền xuyên màng và miền đuôi bào thể (TM/CT). Polypeptit có thể là HA được cải biến và mỗi trong số hai hoặc nhiều hơn hai trình tự axit amin mà tạo thành polypeptit có thể thu được từ các HA khác nhau để tạo ra HA khảm, HA cúm khảm, HA được cải biến khảm hoặc HA cúm

khảm được cải biến. HA khảm cũng có thể bao gồm trình tự axit amin bao gồm peptit tín hiệu khác loại (HA preprotein khảm) mà được tách sau hoặc trong quá trình tổng hợp protein. Tốt hơn là, polypeptit khảm hoặc HA cúm khảm là không xuất hiện tự nhiên. Axit nucleic mã hoá polypeptit khảm có thể được mô tả như "axit nucleic khảm" hoặc "trình tự nucleotit khảm". Ví dụ, axit nucleic khảm có thể bao gồm trình tự nucleotit mã hoá HA được cải biến bao gồm trình tự nucleotit khảm mã hoá, theo dãy, HA ectodomain được cải biến bao gồm vòng phân giải protein được cải biến, miền xuyên màng cúm, và đuôi tế bào chất, trong đó HA ectodomain được cải biến là từ chủng cúm và miền xuyên màng và đuôi tế bào chất là từ chủng cúm thứ hai. Các ví dụ về axit nucleotit khảm, trong đó HA ectodomain được cải biến là từ chủng cúm thứ nhất và miền xuyên màng và đuôi tế bào chất là từ chủng cúm thứ hai được nêu trong các ví dụ 5.14, 5.16, 5.18, 5.19, 5. 21 và 5.23. Hạt tương tự virut gồm có HA khảm có thể được mô tả như "VLP khảm".

Như được mô tả trên đây, protein khảm, polypeptit khảm hoặc HA khảm có thể bao gồm peptit tín hiệu mà là giống như hoặc khác với, phần còn lại của polypeptit hoặc protein. Thuật ngữ "peptit tín hiệu" đã được biết đến trong lĩnh vực này và thường đề cập đến trình tự axit amin ngắn (khoảng từ 5 đến 30 axit amin), thường được tìm thấy ở đầu cuối N của polypeptit mà có thể hoán vị trực tiếp polypeptit dịch mã mới của bào quan cụ thể hoặc trợi giúp việc định vị miền đặc hiệu của chuỗi polypeptit so với miền khác. Đối với ví dụ không hạn chế, peptit tín hiệu có thể tạo đích sự hoán vị của protein vào trong mạng lưới nội chất và/hoặc trợ giúp việc định vị của miền đầu gần đầu cuối N so với miền neo màng của polypeptit mới sinh để trợ giúp cho việc tách và gấp nếp của protein chín, ví dụ HA được cải biến hoặc HA được cải biến khảm.

HA cũng có thể là HA khảm hoặc HA được cải biến khảm, trong đó miền xuyên màng nguyên thể của HA hoặc HA được cải biến là được thay thế bằng miền xuyên màng khác loại. Miền xuyên màng của protein HA được bảo tồn ở mức cao (xem ví dụ Fig.1C của WO 2010/148511). Miền xuyên màng khác loại có thể thu được từ miền xuyên màng HA bất kỳ, ví dụ nhưng không chỉ giới hạn ở miền xuyên màng từ H1 California, B/Florida/4/2006 (Số truy cập GenBank ACA33493.1), B/Malaysia/2506/2004 (Số truy cập GenBank ABU99194.1), H1/Bri (Số truy cập GenBank ADE28750.1), H1

A/Solomon Islands/3/2006 (Số truy cập GenBank ABU99109.1), H1/NC (Số truy cập GenBank AAP34324.1), H2 A/Singapore/1/1957 (Số truy cập GenBank AAA64366.1), H3 A/Brisbane/10/2007 (Số truy cập GenBank ACI26318.1), H3 A/Wisconsin/67/2005 (Số truy cập GenBank ABO37599.1), H5 A/Anhui/1/2005 (Số truy cập GenBank ABD28180.1), H5 A/Vietnam/1194/2004 (Số truy cập GenBank ACR48874.1), H5-Indo (Số truy cập GenBank ABW06108.1). Miền chyển qua màng cũng có thể được xác định bởi trình tự axit amin liên ứng sau đây:

iLXiYystvAiSslXlXXmlagXsXwmcs (SEQ ID NO:94)

Các ví dụ về cấu trúc bao gồm HA khảm với miền xuyên màng khác loại bao gồm: cấu trúc số 1875 (CPMV-HT+ B Brisbane/60/08 với vòng phân giải protein được xóa bỏ + H1TM, với miền xuyên màng và đuôi tế bào chất được thay thế bằng H1 A/California/07/2009; xem ví dụ 5.19), cấu trúc số 1977 (CPMV-160+ B Brisbane/60/08 với vòng phân giải protein được xóa bỏ + H1TM, với miền xuyên màng và đuôi tế bào chất được thay thế bằng H1 A/California/07/2009; xem ví dụ 5.14), cấu trúc số 1067 (CPMV-HT B Brisbane/60/08 với vòng phân giải protein được xóa bỏ + H1TM, với miền xuyên màng và đuôi tế bào chất được thay thế bằng H1 A/California/07/2009; xem ví dụ 5.14), cấu trúc số 2074 (CPMV HT B Massachusetts/2/2012+H1Tm, với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1 A/California/07/2009; xem ví dụ 5.16), cấu trúc số 2060 (CPMV HT160+ Massachusetts/2/2012+H1Tm, với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1 A/California/07/2009; xem ví dụ 5.16), cấu trúc số 2062 (CPMV 160+ B Massachusetts/2/2012+H1Tm, với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1 A/California/07/2009; xem ví dụ 5.21), cấu trúc số 1860 (CPMV HT+ B Wisconsin/1/2010+H1Tm với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1 A/California/07/2009; xem ví dụ 5.23), cấu trúc số 1454 (CPMV HT B Wisconsin/1/2010+H1Tm với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1 A/California/07/2009, xem ví dụ 5.18) và cấu trúc số 1893 (CPMV 160+ B Wisconsin/1/2010+H1Tm với miền xuyên màng và đuôi tế bào chất được thay thế bằng các vòng H1 A/California/07/2009)

xem ví dụ 5.18. Hoạt tính của các HA cải biến khảm này được thể hiện trên các hình vẽ Fig.26B và Fig.27B.

Peptit tín hiệu

Peptit tín hiệu (SP) có thể là nguyên gốc so với HA cải biến hoặc HA cải biến khảm hoặc peptit tín hiệu có thể là khác loại đối với trình tự ban đầu của HA cải biến là được mong đợi. HA cải biến có thể bao gồm peptit tín hiệu từ typ cúm thứ nhất, typ phụ hoặc chủng có sự cân bằng của HA từ một hoặc nhiều hơn một typ cúm khác nhau, typ phụ hoặc chủng. Ví dụ peptit tín hiệu nguyên thể của typ phụ HA H1, H2, H3, H5, H6, H7, H9 hoặc cúm typ B có thể được sử dụng để biểu hiện HA cải biến ở hệ thực vật. Theo một số phương án của sáng chế, SP có thể thuộc cúm typ B, H1, H3 hoặc H5; hoặc typ phụ H1/Bri, H1/NC, H5/Indo, H3/Bri hoặc B/Flo.

Hơn nữa, HA cải biến hoặc HA cải biến khảm có thể bao gồm peptit tín hiệu nguyên thể hoặc không nguyên thể; peptit tín hiệu không nguyên thể có thể có nguồn gốc thực vật hoặc thu được từ polypeptit động vật hoặc vi khuẩn. Peptit tín hiệu nguyên thể có thể tương ứng với peptit tín hiệu của HA hoặc HA cải biến được biểu hiện, ngoài ra, peptit tín hiệu có thể là từ protein cấu trúc hoặc hemagglutinin của virut khác với cúm. Các ví dụ không hạn chế về peptit tín hiệu mà có thể được sử dụng là ví dụ về alfalfa protein disulfua isomeraza (PDI SP; nucleotit từ 32 đến 103 có số truy cập Z11499 cũng xem WO 2009/076778; WO 2010/148511, hoặc WO 2010/003235), hoặc peptit tín hiệu patatin (PatA SP; đị vị nucleotit từ 1738 đến 1806 có Số truy cập Genbank A08215). Trình tự nucleotit của PatA SP đối với số truy cập này là:

ATGGCAACTACTAAAACTTTTTTAATTTTATTTTTTATGATATTAGCAACTACTA
GTTCAACATGTGCT (SEQ ID NO: 171)

trình tự axit amin của peptit tín hiệu patatin A là:

MATTKTFLILFFMILATTSSTCA (SEQ ID NO: 172)

Do đó, sáng chế đề xuất đối với HA cải biến hoặc HA cải biến khảm bao gồm peptit tín hiệu nguyên gốc hoặc không nguyên gốc và axit nucleic mã hoá protein HA cải biến khảm.

Đồng biểu hiện với protein kênh

HA đột biến (được cải biến) có thể được tạo ra ở thực vật bằng cách đồng biểu hiện axit nucleic thứ nhất mã hoá HA cải biến với axit nucleic thứ hai mã hoá protein kênh, ví dụ nhưng không chỉ giới hạn ở protein kênh proton. Axit nucleic thứ nhất và thứ hai có thể được đưa vào thực vật ở cùng bước hoặc chúng có thể được đưa vào thực vật theo trình tự. Axit nucleic thứ nhất và thứ hai có thể được đưa vào thực vật theo cách tạm thời hoặc theo cách ổn định. Hơn nữa, thực vật mà biểu hiện axit nucleic thứ nhất mã hoá HA cải biến có thể được biến nạp với protein kênh, ví dụ nhưng không chỉ giới hạn ở protein kênh proton, (axit nucleic thứ hai) sao cho axit nucleic thứ nhất và thứ hai được đồng biểu hiện ở thực vật. Theo cách khác, thực vật mà biểu hiện protein kênh, ví dụ nhưng không chỉ giới hạn ở protein kênh proton, (axit nucleic thứ hai) có thể được biến nạp với axit nucleic thứ nhất mã hoá HA cải biến sao cho axit nucleic thứ nhất và thứ hai được đồng biểu hiện ở thực vật. Ngoài ra, thực vật thứ nhất biểu hiện axit nucleic thứ nhất mã hoá HA cải biến, có thể được lai ới thực vật thứ hai biểu hiện axit nucleic thứ hai mã hoá protein kênh ví dụ nhưng không chỉ giới hạn ở protein kênh proton, để tạo ra thực vật thế hệ mà đồng biểu hiện axit nucleic thứ nhất và thứ hai mã hoá HA cải biến và protein kênh, ví dụ nhưng không chỉ giới hạn ở protein kênh proton, một cách tương ứng.

Không muốn bị ràng buộc bởi lý thuyết, độ pH của khoang tế bào bao gồm HA cải biến, bao gồm bộ máy golgi, có thể là quan trọng để gấp nếp, ổn định và/hoặc phân giải protein HA. Protein kênh proton, như ví dụ cúm M2 và BM2 protein mà điều hoà độ pH trong khoang tế bào. Ví dụ, M2 điều hoà khả năng dung hợp màng bằng cách tạo đệm khoang nội bào ở giai đoạn trễ và sớm của việc sao chép virut cúm.

Bằng cách đồng biểu hiện protein kênh, ví dụ nhưng không chỉ giới hạn ở protein kênh proton, cùng với HA cải biến, độ pH trong bộ máy golgi có thể gia tăng và dẫn đến

độ ổn định, giảm sự thoái biến hoặc tổ hợp của chúng, và gia tăng mức biểu hiện và sản lượng của HA cải biến và/hoặc VLP.

Bằng cách đồng biểu hiện HA cải biến cùng với protein kênh, ví dụ nhưng không chỉ giới hạn ở protein kênh proton, ở thực vật, sản lượng gia tăng của HA và/hoặc VLP được quan sát, khi so với thực vật mà biểu hiện được cải biến mà không cần phải có sự đồng biểu hiện của protein kênh, ví dụ nhưng không chỉ giới hạn ở protein kênh proton (xem các hình vẽ Fig.13A và Fig.14). Như được thể hiện ví dụ trên Fig.13A, sự đồng biểu hiện của M2 với cúm B HA cải biến làm gia tăng mức kết tụ HA (Fig.13A, 1059 đối lại 1059+1261).

Hơn nữa, hiệu quả của M2 từ cúm A/Puerto Rico/8/1934 làm tăng sự kết tụ của cúm B HA cải biến và H3 so với sự kết tụ của M2 từ cúm A/New Caledonia/20/1999. Đối với cúm B HA cải biến, phép so sánh được thực hiện bằng phép phân tích thấm tách western của phần chiết protein từ thực vật được biến nạp với các cấu trúc 1059, 1059+1261 và 1059+859. Các kết quả thu được chứng minh rằng sự đồng biểu hiện của M2 từ cúm A/Puerto Rico/8/1934 (được mã hoá bởi cấu trúc số 859) là hữu hiệu như sự đồng biểu hiện của M2 từ cúm A/New Caledonia/20/1999 (được mã hoá bởi cấu trúc số 1261) để làm gia tăng sự kết tụ của cúm B HA được cải biến (Fig.14).

Như được sử dụng trong bản mô tả này, thuật ngữ "M2," "M2 protein," "trình tự M2" và "miền M2" đề cập đến tất cả hoặc một phần của trình tự M2 protein được phân lập từ, dựa vào hoặc có mặt trong chủng hoặc thể phân lập virut cúm xuất hiện tự nhiên bất kỳ hoặc được tạo ra về mặt nhân tạo. Do đó, thuật ngữ M2 và tương tự bao gồm các biến thể trình tự M2 xuất hiện tự nhiên được tạo ra bằng việc đột biến trong chu kỳ sống của virut hoặc được tạo ra đáp ứng với áp suất lựa chọn (ví dụ, liệu pháp dược chất, sự mở rộng tính hướng kích thích của tế bào vật chủ hoặc khả năng lây nhiễm, v.v), cũng như trình tự M2 được tạo ra về mặt tái tổ hợp hoặc về mặt tổng hợp. Ví dụ không hạn chế của các trình tự mà có thể được sử dụng với sáng chế bao gồm M2 từ A/Puerto Rico/8/1934 và M2 từ A/New Caledonia/20/1999.

Đáp ứng miễn dịch

"Đáp ứng miễn dịch" thường đề cập đến sự đáp ứng của hệ miễn dịch thích ứng. Hệ miễn dịch thích ứng thường bao gồm đáp ứng thể dịch và đáp ứng qua trung gian tế bào. Đáp ứng thể dịch là khía cạnh của tính miễn dịch mà qua trung gian bởi các kháng thể được tiết, được tạo ra trong tế bào của dòng tế bào bạch huyết B (tế bào B). Các kháng thể được tiết gắn kết với kháng nguyên trên bề mặt của vi sinh vật xâm lấn (như virut hoặc vi khuẩn), mà chỉ báo chúng để phá huỷ. Tính miễn dịch thể dịch được sử dụng thường đề cập đến việc tạo ra kháng thể và quy trình mà kèm theo nó, cũng như chức năng tác quan của kháng thể, bao gồm sự hoạt hoá tế bào Th2 và sự sản xuất xytokin, tạo ra tế bào nhớ, thúc đẩy opsonin của sự thực bào, loại bỏ mầm bệnh và tương tự. Thuật ngữ "điều biến" hoặc "sự điều biến" hoặc tương tự đề cập đến việc tăng hoặc giảm đáp ứng hoặc tham số cụ thể, như được xác định bởi bất kỳ trong số vài thử nghiệm thường đã được biết đến hoặc được sử dụng, một vài trong số chúng được minh họa trong bản mô tả này.

Đáp ứng qua trung gian tế bào là đáp ứng miễn dịch mà không bao gồm các thể nhưng tốt hơn là bao gồm sự hoạt hoá đại thực bào, tế bào giết tự nhiên (NK), tế bào bạch huyết T độc tế bào đặc hiệu kháng nguyên và giải phóng các xytokin khác nhau trong sự đáp ứng với kháng nguyên. Tính miễn dịch qua trung gian tế bào thường được sử dụng đề cập đến một vài sự hoạt hoá tế bào Th, sự hoạt hoá tế bào Tc và sự đáp ứng qua trung gian tế bào T. Tính miễn dịch qua trung gian tế bào là quan trọng đặc biệt trong đáp ứng với sự lây nhiễm virut.

Ví dụ, việc cảm ứng tế bào bạch huyết T dương tính CD8 đặc hiệu kháng nguyên có thể được đo bằng cách sử dụng thử nghiệm ELISPOT; sự kích thích tế bào bạch huyết T dương tính CD4 có thể được đo bằng cách sử dụng thử nghiệm tăng sinh tế bào. Độ chuẩn kháng thể kháng cúm có thể được định lượng bằng cách sử dụng thử nghiệm ELISA; lớp kháng thể phản ứng đặc hiệu kháng nguyên hoặc phản ứng ngang cũng có thể được đo bằng cách sử dụng kháng thể kháng lớp kháng nguyên (ví dụ kháng IgG, IgA, IgE hoặc IgM). Các phương pháp và kỹ thuật để thực hiện các thử nghiệm này đã được biết đến trong lĩnh vực này.

Độ chuẩn HAI phản ứng ngang cũng có thể được sử dụng để chứng minh hiệu quả của đáp ứng miễn dịch với các chủng virut khác đề cập đến vacxin typ phụ. Ví dụ, huyết thanh từ đối tượng được gây miễn dịch bằng chế phẩm vacxin chứa chủng thứ nhất (ví dụ VLP của A/Indonesia 5/05) có thể được sử dụng trong thử nghiệm HAI với chủng thứ hai của virut nguyên vẹn hoặc hạt virut (ví dụ A/Vietnam/1194/2004), và độ chuẩn HAI được xác định.

Sự có mặt hoặc nồng độ xytokin cũng có thể được định lượng. Ví dụ, sự đáp ứng tế bào hỗ trợ T (Th1/Th2) sẽ được đặc trưng bằng việc đo tế bào tiết IFN-γ và IL-4 bằng cách sử dụng bởi ELISA (ví dụ BD Biosciences OptEIA kits). Tế bào đơn nhân máu ngoại vi (PBMC) hoặc tế bào lá lách thu được từ đối tượng có thể được nuôi cấy và dịch nổi được phân tích. Tế bào bạch huyết T cũng có thể được định lượng bằng việc sắp xếp tế bào hoạt hoá huỳnh quang (FACS), bằng cách sử dụng chất đánh dấu huỳnh quang đặc hiệu đánh dấu và phương pháp là đã được biết đến trong lĩnh vực này.

Thử nghiệm vi trung hoà cũng có thể được thực hiện để mô tả đặc điểm đáp ứng miễn dịch ở đối tượng, xem ví dụ các phương pháp của ấn phẩm: Rowe et al., 1973. Các độ chuẩn trung hoà virut có thể thu được bằng vài cách, bao gồm: 1) đếm vết tan tiêu (thử nghiệm vết tan) sau khi cố định/tạo màu tím tinh thể của tế bào; 2) quan sát bằng kính hiển vi dịch tan tế bào trong môi trường nuôi cấy; 3) ELISA và phát hiện quang phổ của protein virut NP (tương quan với sự lây nhiễm virut của tế bào chủ).

Thuật ngữ "hạt tương tự virut" (VLP), hoặc "các hạt tương tự virut" hoặc "VLP" đề cập đến các cấu trúc mà tự lắp ráp và bao gồm protein virut ví dụ cúm protein HA hoặc protein HA cải biến như ví dụ HA0 protein, trong đó vòng phân giải protein đã được cải biến. VLP thường tương tự về mặt hình thái học và về mặt kháng nguyên với virion được tạo ra khi lây nhiễm, nhưng thiếu thông tin di truyền đủ để sao chép và do đó là không lây nhiễm. Trong một số ví dụ, VLP có thể bao gồm các loại protein đơn lẻ hoặc nhiều hơn một loại protein. Đối với VLP bao gồm nhiều hơn một loại protein, các loại protein có thể là từ cùng loài virut hoặc có thể bao gồm protein từ các loài khác nhau, chi, liên họ hoặc họ virut (như được xác định bởi danh pháp ICTV). Trong các ví dụ khác, một hoặc nhiều loại protein bao gồm VLP có thể được cải biến từ trình tự xuất hiện tự nhiên, như ví dụ

HA cải biến như được mô tả trong bản mô tả này. VLP có thể được tạo ra ở tế bào vật chủ thích hợp bao gồm tế bào vật chủ thực vật và côn trùng. Sau khi chiết từ tế bào vật chủ và khi phân lập và tinh chế hơn nữa trong các điều kiện thích hợp, VLP có thể được tinh chế như các cấu trúc nguyên vẹn.

Hơn nữa, VLP có thể được tạo ra mà bao gồm tổ hợp của typ phụ HA. Ví dụ, VLP có thể bao gồm một hoặc nhiều hơn một HA hoặc một hoặc nhiều hơn một HA cải biến từ typ phụ H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, typ phụ B HA hoặc tổ hợp của chúng. Việc lựa chọn tổ hợp của HA hoặc HA cải biến có thể được xác định bằng việc sử dụng định trước vacxin được tạo ra từ VLP. Ví dụ, vacxin để sử dụng trong việc tiêm truyền chim có thể bao gồm tổ hợp bất kỳ của typ phụ HA hoặc typ phụ HA được cải biến, trong khi VLP hữu ích để tiêm truyền cho con người có thể bao gồm typ phụ một hoặc nhiều hơn một typ phụ hoặc typ phụ cải biến của H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15, H16, typ phụ B HA. Tuy nhiên, tổ hợp typ phụ HA khác hoặc typ phụ HA cải biến có thể được tạo ra phụ thuộc vào việc sử dụng VLP. Để tạo ra VLP bao gồm tổ hợp của typ phụ HA hoặc typ phụ HA cải biến, typ phụ HA mong muốn hoặc typ phụ HA cải biến có thể được đồng biểu hiện trong cùng tế bào, ví dụ tế bào thực vật.

VLP được tạo ra từ protein có nguồn gốc cúm, theo sáng chế không bao gồm M1 protein. M1 protein đã được biết để gắn kết ARN (Wakefield and Brownlee, 1989) mà là chất gây ô nhiễm của chế phẩm VLP. Sự có mặt của ARN là mong muốn khi thu được sự chấp thuận điều hoà đối với sản phẩm VLP, do đó chế phẩm VLP thiếu ARN có thể là có lợi.

VLP được tạo ra như được mô tả trong bản mô tả này thường không bao gồm neuramindaza (NA). Tuy nhiên, NA có thể được đồng biểu hiện với HA nếu VLP bao gồm HA và NA được mong muốn.

Sáng chế cũng bao gồm, nhưng không chỉ giới hạn ở, VLP thu được từ virut mà thu được vỏ lipit từ màng sinh chất của tế bào trong đó VLP protein được biểu hiện. Ví dụ,

nếu VLP được biểu hiện ở hệ dựa trên thực vật, VLP có thể thu được vỏ lipit từ màng sinh chất.

Thông thường, thuật ngữ "lipit" đề cập đến các phân tử cuất hiện tự nhiên có thể hoà tan trong chất béo (ưa lipit). Thuật ngữ này cũng được sử dụng cụ thể hơn đề cập đến các axit béo và dẫn xuất của chúng (bao gồm tri-, di-, và monoglyxerit và phospholipit), cũng như chất chuyển hoá chứa sterol có thể hoà tan trong chất béo khác hoặc sterol. Phospholipit là thành phần chính của tất cả cá màng sinh học, cùng với glycolipit, sterol và protein. Các ví dụ về phospholipit bao gồm phosphatidyletanolamin, phosphatidylcholin, phosphatidylinositol, phosphatidylserin và tương tự. Các ví dụ về sterol bao gồm zoosterol (ví dụ, cholesterol) và phytosterol. Hơn 200 phytosterol đã được nhận dạng trong các loài thực vật khác nhau, phổ biến nhất là campesterol, stigmasterol, ergosterol, brassicasterol, delta-7-stigmasterol, delta-7-avenasterol, daunosterol, sitosterol, 24-metylcholesterol, cholesterol hoặc beta-sitosterol. Chuyên gia trong lĩnh vực này sẽ hiểu được, chế phẩm lipit chứa màng sinh chất có thể thay đổi với môi trường nuôi cấy hoặc điều kiện sinh trưởng của tế bào hoặc sinh vật mà từ đó thu được tế bào.

Màng sinh chất thường bao gồm hai lớp lipit, cũng như protein đối với các chức năng khác nhau. Nồng độ khu trú của lipit cụ thể có thể được tìm thấy trong hai lớp lipit, đề cập đến 'mảng lipit'. Không muốn bị ràng buộc bởi lý thuyết, mảng lipit có thể có vai trò đáng kể trong nội độc tố và ngoại độc tố, đi vào hoặc đi ra của virut hoặc các chất lây nhiễm khác, truyền tín hiệu nội bào, tương tác với các thành phần cấu trúc khác của tế bào hoặc sinh vật, như nền nội bào và ngoại bào.

Ở thực vật, cúm VLP mọc chồi từ màng sinh chất do đó chế phẩm lipit của VLP phản ánh nguồn gốc của chúng. VLP được tạo ra theo sáng chế bao gồm HA của một hoặc nhiều hơn một typ hoặc typ phụ cúm, tạo phức hệ với lipit thu được từ thực vật. Lipit thực vật có thể kích thích tế bào miễn dịch đặc hiệu và làm tăng sự đáp ứng miễn dịch được gây ra. Màng thực vật được tạo thành từ lipit, phosphatidylcholin (PC) và phosphatidyletanolamin (PE), và cũng chứa glycosphingolipit, saponin và phytosterol. Ngoài ra, mảng lipit cũng được tìm thấy ở màng sinh chất thực vật – các vi miền này được

làm giàu trong sphingolipit và sterol. Ở thực vật, nhiều phytosterol đã biết xuất hiện, bao gồm stigmasterol, sitosterol, 24-metylcholesterol và cholesterol (Mongrand et al., 2004).

PC và PE, cũng như glycosphingolipit có thể gắn kết với phân tử CD1 được biểu hiện bởi tế bào miễn dịch động vật có vú như tế bào trình diện kháng nguyên (APC) tương tự tế bào hình cây và đại thực bào và các tế bào bao gồm tế bào bạch huyết B và T trong tuyến ước và gan (Tsuji M,. 2006). Các phân tử CD1 tương tự về mặt cấu trúc với phần lớn phân tử phức hệ tương thích ghép mô (MHC) thuộc nhóm I và vai trò của chúng thể hiện các kháng nguyên glycolipit với tế bào NKT (tế bào T giết tự nhiên). Khi hoạt hoá, tế bào NKT hoạt hoá tế bào miễn dịch di truyền như tế bào NK và tế bào hình cây và cũng hoạt hoá tế bào miễn dịch thích ứng tương tự tế bào B và tế bào T tạo ra kháng thể.

Nhiều phytosterol có thể được tìm thấy trong màng sinh chất-thành phần đặc hiệu có thể thay đổi phụ thuộc vào các loài, điều kiện sinh trưởng, nguồn dinh dưỡng hoặc trạng thái mầm bệnh, để đặt tên một vài nhân tố. Thông thường, beta-sitosterol là phytosterol nhiều nhất.

Phytosterol có mặt ở cúm VLP được tạo phức hệ với hai lớp lipit, như vỏ thu được từ màng sinh chất có thể tạo ra chế phẩm vacxin có lợi. Không muốn bị ràng buộc bởi lý thuyết, VLP được tạo ra từ thực vật được tạo phức với hai lớp lipit, như vỏ thu được từ mang sinh chất, có thể gây ra phản ứng miễn dịch mạnh hơn so với VLP được tạo ra trong hệ biểu hiện khác và có thể tương tự với phản ứng miễn dịch do vacxin virut sống hoặc nguyên vẹn giảm độc lực gây ra.

VLP như được mô tả trong bản mô tả này có thể được tạo phức hệ với hai lớp lipit thu được từ thực vật. Theo một số phương án, hai lớp lipit thu được từ thực vật có thể bao gồm vỏ của VLP. Lipit thu được từ thực vật có thể bao gồm các thành phần lipit của màng sinh chất của thực vật trong đó VLP được tạo ra, bao gồm, nhưng không chỉ giới hạn ở, phosphatidylcholin (PC), phosphatidyletanolamin (PE), glycosphingolipit, phytosterol hoặc tổ hợp của chúng. Theo cách khác, lipit thu được từ thực vật có thể được đề cập đến 'lipit thực vật'. Các ví dụ về phytosterol là đã được biết đến trong lĩnh vực này và bao

gồm, ví dụ, stigmasterol, sitosterol, 24-metylcholesterol và cholesterol – xem, ví dụ, ấn phẩm: Mongrand et al., 2004.

VLP có thể được đánh giá đối với cấu trúc và kích cỡ bằng, ví dụ, thử nghiệm hemagglutin hoá, kính hiển vi điện tử hoặc bằng việc sắc ký ngoại cỡ.

Đối với sắc ký ngoại cỡ, tổng protein có thể hoà tan có thể được chiết từ mô thực vật bằng cách làm đồng nhất (Polytron) mẫu của vật liệu thực vật nghiền đông lạnh trong chất đệm chiết và vật liệu không thể hoà tan được loại bỏ bằng việc ly tâm. Việc kết tủa với PEG có thể được sử dụng. Protein có thể hoà tan được định lượng và phần chiết đi qua nền ngoại cỡ, ví dụ nhưng không chỉ giới hạn ở Sephacryl™. Sau khi sắc ký, các phân đoạn còn có thể được phân tích bằng việc thấm tách miễn dịch để xác định phần bù protein của phân đoạn.

Không muốn bị ràng buộc bởi lý thuyết, khả năng của HA gắn kết với RBC từ các động vật khác nhau được điều khiển bằng ái lực HA đối với axit sialic α2,3 hoặc α2,3 và sự có mặt của các axit sialic này trên bề mặt của RBC. HA ngực và chim từ virut cúm kết dính hồng cầu từ tất cả vài loài bao gồm gà tây, gà, vịt, chuột lang, con người, cừu, ngựa và bò; trong khi HA của người sẽ gắn kết với hồng cầu của gà tây, gà, vịt, chuột lang, con người và cừu (cũng xem ấn phẩm: Ito T. et al, 1997, Virology, vol 227, p493-499; và Medeiros R et al, 2001, Virology, vol 289 p.74-85).

Việc gấp đúng chính xác của protein virut được biểu hiện có thể là quan trọng đối với tính ổn định của protein, sự hình thành multime, sự hình thành VLP, chức năng của protein virut và sự nhận biết protein virut bởi kháng thể, trong số các đặc điểm khác. Việc gấp nếp và kết tụ của protein có thể được tác động bởi một hoặc nhiều yếu tố, bao gồm, nhưng không chỉ giới hạn ở, trình tự protein, lượng dư tương đối của protein, mức độ tập hợp nội bào, độ pH trong khoang tế bào, mức sinh khả dụng của đồng yếu tố mà có thể gắn kết hoặc được kết hợp theo cách tạm thời với protein được gấp nếp, gấp nếp riêng phần hoặc không gấp nếp, sự có mặt của một hoặc nhiều chaperone protein hoặc tương tự.

Protein sốc nhiệt (Hsp) hoặc protein ứng suất là các ví dụ về chaperon protein, mà có thể tham gia vào các quy trình tế bào khác nhau bao gồm việc tổng hợp protein, nghẽn

nội bào, ngăn ngừa việc gấp nếp lỗi, ngăn ngừa sự kết tụ protein, lắp ráp và tháo rời phức hệ protein, gấp nếp protein và phân rã protein. Các ví dụ về chaperon protein bao gồm, nhưng không chỉ giới hạn ở, Hsp60, Hsp65, Hsp 70, Hsp90, Hsp100, Hsp20-30, Hsp10, Hsp100-200, Hsp100, Hsp90, Lon, TF55, FKBPs, xyclophilin, ClpP, GrpE, ubiquitin, calnexin, và protein disulfua isomeraza (xem, ví dụ, Macario, A.J.L., *Cold Spring Harbor Laboratory Res.* 25:59-70. 1995; Parsell, D.A. & Lindquist, S. *Ann. Rev. Genet.* 27:437-496 (1993); patent Mỹ số 5,232,833). Như được mô tả trong bản mô tả này, chaperon protein, ví dụ nhưng không chỉ giới hạn ở Hsp40 và Hsp70 có thể được sử dụng để bảo đảm việc gấp nếp của protein virut.

Các ví dụ về Hsp70 bao gồm Hsp72 và Hsc73 từ tế bào động vật có vú, DnaK từ vi khuẩn, cụ thể là mycobacteria như *Mycobacterium* leprae, *Mycobacterium tuberculosis, và Mycobacterium bovis* (như Bacille-Calmette Guerin: được đề cập trong bản mô tả này là Hsp7l). DnaK từ *Escherichia coli,* nấm men và các sinh vật nhân sơ khác và BiP và Grp78 từ sinh vật nhân chuẩn như *A. thaliana* (Lin et al. 2001 (Cell Stress and Chaperons 6:201-208). Ví dụ cụ thể về Hsp70 là *A. thaliana* Hsp70 (được mã hoá bởi Genbank ref: AY120747.1). Hsp70 có khả năng gắn kết một cách đặc hiệu ATP cũng như polypeptit không gấp nếp và peptit, do đó tham gia vào việc gấp nếp và không gấp nếp protein cũng như trong việc lắp ráp và tháo rời phức hệ protein.

Các ví dụ về Hsp40 bao gồm DnaJ từ các sinh vật nhân sơ như *E. coli và* mycobacteria và HSJ1, HDJl và Hsp40 từ sinh vật nhân chuẩn, như alfalfa (Frugis et al., 1999. Plant Molecular Biology 40:397-408). Ví dụ cụ thể về Hsp40 là *M. sativa* MsJ1 (Genbank ref: AJ000995.1). Hsp40 đóng vai trò như chaperon phân tử trong việc gấp nếp protein, sự dung nạp nhiệt và sao chép ADN, trong số các hoạt tính tế bào khác.

Trong số Hsps, Hsp70 và đồng chaperon của nó, Hsp40, có liên quan đến sự ổn định hoá việc dịch mã và polypeptit được tổng hợp mới trước khi việc tổng hợp hoàn thiện. Không muốn bị ràng buộc bởi lý thuyết, Hsp40 gắn kết với miếng đắp kỵ nước của polypeptit không gấp nếp (mới sinh hoặc được lây truyền mới), do đó tạo thuận lợi cho sự tương tác của phức hệ Hsp70-ATP với polypeptit. Sự thuỷ phân ATP dẫn đến sự hình thành của phức hệ ổn định giữa polypeptit, Hsp70 và ADP, và giải phóng Hsp40. Sự kết

hợp của phức hệ Hsp70-ADP với miếng đắp kỵ nước của polypeptit ngăn ngừa sự tương tác của chúng với các miếng đắp kỵ nước khác, ngăn ngừa sự gấp nếp không đúng và sự hình thành khối kết tụ với các protein khác (xem trong ấn phẩm: Hartl, FU. 1996. Nature 381:571-579).

Chaperon protein nguyên thể có thể tạo thuận lợi cho việc gấp nếp đúng có hàm lượng protein tái tổ hợp thấp, nhưng mức biểu hiện gia tăng, sự dư thừa của chaperon nguyên gốc có thể trở thành nhân tố giới hạn. Mức biểu hiện cao của protein virut trong lá thấm lọc nông nghiệp có thể dẫn đến sự kết tụ của protein virut trong xytosol, và sự đồng biểu hiện của một hoặc nhiều hơn một chaperon protein như Hsp70, Hsp40 hoặc Hsp70 và Hsp40 có thể làm giảm hàm lượng protein gấp nếp lỗi hoặc kết tụ và gia tăng số lượng protein thể hiện các đặc điểm cấu trúc bậc ba và bậc bốn mà cho phép sự hình thành của hạt tương tự virut.

Do đó, sáng chế cũng đề xuất phương pháp tạo ra protein virut VLP ở thực vật, trong đó axit nucleic thứ nhất mã hoá protein virut được đồng biểu hiện với axit nucleic thứ hai mã hoá protein kênh, ví dụ nhưng không chỉ giới hạn ở protein kênh proton, và axit nucleic thứ ba mã hoá chaperon. Axit nucleic thứ nhất, thứ hai và thứ ba có thể được đưa vào thực vật theo cùng bước hoặc có thể được đưa vào thực vật theo trình tự.

VLP được tạo ra trong thực vật có thể cảm ứng protein virut bao gồm N-glycan đặc hiệu thực vật. Do đó, sáng chế cũng đề xuất VLP bao gồm protein virut có N-glycan đặc hiệu thực vật.

Hơn nữa, sự cải biến của N-glycan ở thực vật là đã được biết đến (xem ví dụ WO 2008/151440; WO 2010/006452; hoặc U.S. 60/944,344) và protein virut có N-glycan cải biến có thể được tạo ra. Protein virut bao gồm mẫu glycosyl hoá cải biến, ví dụ với N-glycan được fucosyl hoá, xylosyl hoá hoặc cả hai, fucosyl hoá và xylosyl hoá, có thể thu được hoặc protein virut có mẫu glycosyl hoá cải biến có thể thu được, trong đó protein thiếu sự fucosyl hoá, xylosyl hoá hoặc cả hai và bao gồm galatosyl hoá gia tăng. Hơn nữa, việc điều biến của các cải biến sau dịch mã, ví dụ, việc bổ sung galactoza đầu cuối có thể

dẫn đến việc giảm fucosyl hoá và xylosyl hoá của protein virut được biểu hiện khi so với thực vật kiểu hoang dại biểu hiện protein virut.

Ví dụ, mà không được cho là giới hạn, việc tổng hợp protein virtu có mẫu glycosyl hoá cải biến có thể đạt được bằng cách đồng biểu hiện protein quan tâm cùng với trình tự nucleotit mã hoá beta-1.4galactosyltransferaza (GalT), ví dụ, nhưng không chỉ giới hạn ở GalT động vật có vú hoặc GalT của con người, tuy nhiên GalT từ các nguồn khác cũng có thể được sử dụng. Miền xúc tác của GalT cũng có thể được dung hợp với miền CTS (nghĩa là đuôi tế bào chất, miền xuyên màng, vùng gốc) của N-axetylglucosaminyl transferaza (GNT1), để tạo ra enzym lai GNT1-GalT và enzym lai có thể được đồng biểu hiện với protein virut. Protein virut cũng có thể được đồng biểu hiện cùng với trình tự nucleotit mã hoá N-axetylglucosaminyltrasnferaza III (GnT-III), ví dụ nhưng không chỉ giới hạn ở GnT-III của động vật có vú hoặc GnT-III của con người, GnT-III từ các nguồn khác cũng có thể được sử dụng. Ngoài ra, enzym lai GNT1-GnT-III, bao gồm CTS của GNT1 được dung hợp với GnT-III cũng có thể được sử dụng.

Do đó, sáng chế cũng bao gồm VLP bao gồm một hoặc nhiều protein virut có N-glycan cải biến.

Ví dụ không hạn chế về trình tự mà có thể được sử dụng với sáng chế để tạo ra HA cải biến cũng bao gồm các trình tự được mô tả trong các công bố đơn quốc tế số WO 2009/009876; WO 2009/076778; WO 2010/003225; WO 2010/148511; WO 2010/003235; WO 2010/006452, ví dụ, nhưng không chỉ giới hạn ở:

H1 protein được mã hoá bởi phân tử axit nucleic ví dụ từ chủng A/Brisbane/59/2007 (H1N1), A/New Caledonia/20/99 (H1N1), A/Solomon Islands 3/2006 (H1N1), /PuertoRico/8/34 (H1N1), A/Brisbane/59/2007 (H1N1);

H2 protein được mã hoá bởi phân tử axit nucleic có thể là từ chủng A/Singapore/1/57 (H2N2);

H3 protein được mã hoá bởi phân tử axit nucleic có thể là từ chủng A/Brisbane 10/2007 (H3N2), A/Wisconsin/67/2005 (H3N2), A/Victoria/361/2011 (H3N2) hoặc A/Perth/16/2009 (H3N2);

H5 protein được mã hoá bởi phân tử axit nucleic có thể là từ A/Anhui/1/2005 (H5N1), A/Indonesia/5/2005 (H5N1), A/Vietnam/1194/2004 (H5N1)

H6 protein được mã hoá bởi phân tử axit nucleic có thể là từ chủng A/Teal/HongKong/W312/97 (H6N1);

H7 protein được mã hoá bởi phân tử axit nucleic cũng có thể là từ chủng A/ Hangzhou/1/13 (H7N9), A/Equine/Prague/56 (H7N7);

H9 protein được mã hoá bởi phân tử axit nucleic có thể là từ chủng A/HongKong/1073/99 (H9N2);

HA protein từ typ phụ B được mã hoá bởi axit nucleic có thể là từ chủng B/Florida/4/2006, B/Massachusetts/2/12, B/Malaysia/2506/2004, B/Wisconsin/1/2010, hoặc B/Brisbane/60/2008.

Bảng 3: Ví dụ về các cấu trúc mà đã được tạo ra như được mô tả trong bản mô tả này:

Cấu trúc số	Vùng tăng cường sự biểu hiện	Yếu tố khuếch đại	Mô tả	Ví dụ
1261	CPMV-HT		A/New Caledonia/20/1999 (H1N1)	Ví dụ 5.1
859	CPMV-HT		A/Puerto Rico/8/1934 (H1N1)	Ví dụ 5.2
1008	CPMV-HT	BeYDV	B/Brisbane/60/2008	Ví dụ 5.3
1059	CPMV-HT	BeYDV	B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ	Ví dụ 5.4
1462	CPMV-HT	BeYDV	B/Wisconsin/1/2010	Ví dụ 5.5
1467	CPMV-HT	BeYDV	B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ	Ví dụ 5.6
1039	CPMV-HT		B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ	Ví dụ 5.7
1029	CPMV-HT		B/Brisbane/60/2008	Ví dụ 5.11
1829	CPMV-HT+		B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ	Ví dụ 5.12

1937	CPMV-160+		B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ	Ví dụ 5.13
1067	CPMV-HT		B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ +H1 California TMCT	Ví dụ 5.14
1977	CPMV-160+		B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ +H1 California TMCT	Ví dụ 5.14
1875	CPMV-HT+		B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ +H1 California TMCT	Ví dụ 5.19
676	CPMV-HT		H5 A/Indonesia/5/2005 với vị trí tách TETR	Ví dụ 5.8
766	CPMV-HT		H5 A/Indonesia/5/2005 với vị trí tách TETQ	Ví dụ 5.9
928	CPMV-HT		H5 A/Indonesia/5/2005 với vòng phân giải protein được xóa bỏ	Ví dụ 5.10
489	CPMV-HT		H5 A/Indonesia/5/2005 (nguyên gốc)	Ví dụ 5.24
2220	CPMV-HT+		H2 A/Singapore/1/57 (nguyên gốc)	Ví dụ 5.27
2221	CPMV-HT+		H2 A/Singapore/1/57 với vòng phân giải protein được xóa bỏ	Ví dụ 5.28
2222	CPMV-160+		H2 A/Singapore/1/57 (nguyên gốc)	Ví dụ 5.29
2223	CPMV-160+		H2 A/Singapore/1/57 với vòng phân giải protein được xóa bỏ	Ví dụ 5.29
2019	CPMV-HT+		H3 A/Perth/16/09 (nguyên gốc)	Ví dụ 5.30
2039	CPMV-HT+		H3 A/Perth/16/09 với vòng phân giải protein được xóa bỏ	Ví dụ 5.31
2139	CPMV-160+		H3 A/Perth/16/09 (nguyên gốc)	Ví dụ 5.30
2159	CPMV-160+		H3 A/Perth/16/09 với vòng phân giải protein được xóa bỏ	Ví dụ 5.31
1819	CPMV-HT+		H3 A/Victoria /361/11 (nguyên gốc)	Ví dụ 5.26
2230	CPMV-HT+		H3 A/Victoria /361/11 với vòng phân giải protein được xóa bỏ	Ví dụ 5.32
1800	CPMV-160+		H3 A/Victoria /361/11 (nguyên gốc)	Ví dụ 5.25
2250	CPMV-160+		H3 A/Victoria /361/11 với vòng phân giải protein được xóa bỏ	Ví dụ 5.32
2142	CPMV-HT+		H7 A/Hangzhou/1/13 (nguyên gốc)	Ví dụ 5.33
2152	CPMV-HT+		H7 A/Hangzhou/1/13 với vòng phân giải protein được xóa bỏ	Ví dụ 5.34

2224	CPMV-HT+		H9 A/Hong Kong/1073/99 (nguyên gốc)	Ví dụ 5.35
2225	CPMV-HT+		H9 A/Hong Kong/1073/99 với vòng phân giải protein được xóa bỏ	Ví dụ 5.36
2226	CPMV-160+		H9 A/Hong Kong/1073/99 (nguyên gốc)	Ví dụ 5.35
2227	CPMV-160+		H9 A/Hong Kong/1073/99 với vòng phân giải protein được xóa bỏ	Ví dụ 5.36
2013	CPMV-160+		B/Malaysia /2506/04 (nguyên gốc)	Ví dụ 5.37
2014	CPMV-160+		B/Malaysia /2506/04 với vòng phân giải protein được xóa bỏ	Ví dụ 5.38
2070	CPMV-HT		B/Massachusetts/2/12 (nguyên gốc)	Ví dụ 5.39
2072	CPMV-HT		B/Massachusetts/2/12 với vòng phân giải protein được xóa bỏ	Ví dụ 5.15
2080	CPMV-HT+		B/Massachusetts/2/12 (nguyên gốc)	Ví dụ 5.39
2052	CPMV-HT+		B/Massachusetts/2/12 với vòng phân giải protein được xóa bỏ	Ví dụ 5.20
2090	CPMV-160+		B/Massachusetts/2/12 (nguyên gốc)	Ví dụ 5.39
2050	CPMV-160+		B/Massachusetts/2/12 với vòng phân giải protein được xóa bỏ	Ví dụ 5.15
2074	CPMV HT		HA B/Massachusetts/2/12 (PrL-)+H1 California TMCT	Ví dụ 5.16
2060	CPMV-160+		HA B/Massachusetts/2/12 (PrL-)+H1 California TMCT	Ví dụ 5.16
2062	CPMV-HT+		HA B/Massachusetts/2/12 (PrL-)+H1 California TMCT	Ví dụ 5.21
1445	CPMV HT		B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ (PrL-)	Ví dụ 5.17
1839	CPMV-HT+		B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ (PrL-)	Ví dụ 5.22
1820	CPMV160+		B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ (PrL-)	Ví dụ 5.17
1975	CPMV160		B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ (PrL-)	Ví dụ 5.17
1454	CPMV-HT		HA B Wisconsin với vòng phân giải protein được xóa bỏ (PrL-)+H1 California TMCT	Ví dụ 5.18
1893	CPMV-160+		HA B Wisconsin với vòng phân giải protein được xóa bỏ (PrL-)+H1 California TMCT	Ví dụ 5.18
1860	CPMV-HT+		HA B Wisconsin với vòng phân giải protein được xóa bỏ (PrL-)+H1	Ví dụ 5.23

			California TMCT	
2102	CPMV-HT+		B Florida với vòng phân giải protein được xóa bỏ (PrL-)	Ví dụ 5.40
2104	CPMV-HT+	BeYDV	B Florida với vòng phân giải protein được xóa bỏ (PrL-)	Ví dụ 5.40
2016	CPMV-HT+		B Florida +H1 California TMCT với vòng phân giải protein được xóa bỏ (PrL-)	Ví dụ 5.41
2108	CPMV-HT+	BeYDV	B Florida +H1 California TMCT với vòng phân giải protein được xóa bỏ (PrL-)	Ví dụ 5.41

Bảng 4: Mô tả trình tự

SEQ ID NO:	Mô tả	Hình vẽ	SEQ ID NO:	Mô tả	Hình vẽ
1	Trình tự liên ứng vòng phân giải protein H5 của chim		102	Trình tự axit amin của HA B Wisconsin (PrL-).	36B
2	IF-H5A-I-05.si+3c	1A	103	Trình tự nucleotit của HA B Wisconsin (PrL-)+H1 California TMCT	37A
3	IF-H5dTm.r	1B	104	Trình tự axit amin của HA B Wisconsin (PrL-)+H1 California TMC.	37B
4	Cấu trúc 1191	1D	105	Trình tự nucleotit của PDISP/HA B Brisbane (PrL-)+H1 California TMCT.	38A
5	Cátxét 489	1E	106	Trình tự axit amin của PDISP/HA B Brisbane (PrL-)+H1 California TMCT.	38B
6	Trình tự axit amin H5A/Indonesia/5/2005 (H5N1)	1F	107	Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-).	39A
7	IF-S1-M1+M2ANC.c	2A	108	Trình tự axit amin của PDISP/HA B Massachussetts (PrL-).	39B
8	IF-S1-4-M2ANC.r	2B	109	Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-	40A

SEQ ID NO:	Mô tả	Hình vẽ	SEQ ID NO:	Mô tả	Hình vẽ
				)+H1 California TMCT.	
9	M2 tổng hợp (nt 1-26 nối với 715-982 từ DQ508860)	2C	110	Trình tự axit amin của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT.	40B
10	Cátxét 1261	2D	111	Trình tự nucleotit của HA B Wisconsin (PrL-).	41A
11	M2 từ cúm A/New Caledonia/20/1999 (H1N1)	2E	112	Trình tự axit amin của HA B Wisconsin (PrL-).	41B
12	M2 tổng hợp (nt 26-51 nối với nt 740-1007 từ EF467824)	3A	113	Trình tự nucleotit của HA B Wisconsin (PrL-)+H1 California TMCT	42A
13	Cátxét 859	3B	114	Trình tự axit amin của HA B Wisconsin (PrL-)+H1 California TMC.	42B
14	Trình tự axit amin M2 cúm A/Puerto Rico/8/1934 (H1N1)	3C	115	Trình tự nucleotit của H5 Indonesia. nguyên gốc	43A
15	Cátxét 1039	8A	116	Trình tự axit amin của H5 Indonesia nguyên gốc	43B
16	Trình tự axit amin B/Brisbane/60/2008	21A	117	IF**(SacII)-PDI.s1+4c	44A
17	Trình tự axit amin delta-vòng phân giải protein của typ B HA (với chất đánh dấu GG)	21B	118	IF-H3V36111.s1-4r	44B
18	Trình tự axit amin thay thế vị trí tách của B HA với cầu liên kết	21C	119	Trình tự nucleotit của PDISP/H3 Victoria.	44C
19	Trình tự axit amin H3 A/Perth/16/2009	21D	120	Cấu trúc 2171	44E
20	Trình tự axit amin delta-vòng phân giải protein H3 (với cầu liên kết GS)	21E	121	Cấu trúc 1800	44F
21	Trình tự axit amin thay thế vị trí tách của H3 với cầu liên kết	21F	122	Trình tự axit amin của PDISP/H3 Victoria	44G
22	vùng liên kết H1 New Cal	15	123	IF(SacII)-Kozac_PDI.c	45A

SEQ ID NO:	Mô tả	Hình vẽ	SEQ ID NO:	Mô tả	Hình vẽ
23	vùng liên kết H1 Brisbane	15	124	IF-H3V36111.s1-4r	45B
24	vùng liên kết H1 Sol Islands	15	125	Cấu trúc 2181	45D
25	vùng liên kết H2A Singapore	15	126	Cấu trúc 1819	45E
26	vùng liên kết H3A Brisbane	15	127	IF**-H2S157.s1-6r	48A
27	vùng liên kết H3A WCN	15	128	Trình tự nucleotit của PDISP/H2 Singapore.	48B
28	vùng liên kết H5 Anhui	15	129	Cátxét biểu hiện số 2220	48C
29	vùng liên kết H5 Indo	15	130	Trình tự axit amin của PDISP/H2 Singapore	48D
30	vùng liên kết H5 Vietnam	15	131	H2S157(Prl-).r	49A
31	Cấu trúc 1194	4B	132	H2S157(Prl-).c	49B
32	Cátxét 1008	4C	133	Cátxét biểu hiện số 2221	49C
33	vùng liên kết H6 Teal HK	15	134	Trình tự axit amin của PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ	49D
34	vùng liên kết H7 Eq Prague	15	135	Cátxét biểu hiện số 2222	50A
35	vùng liên kết H9A HK	15	136	Cátxét biểu hiện số 2223	50B
36	B Florida vùng liên kết	15	137	Trình tự nucleotit của PDISP/H3 Perth	51A
37	B Malaysia	15	138	IF**-H3P1609.S1-6r	51B
38	1039+1059.r	5A	139	Trình tự axit amin của PDISP/H3 Perth	51C
39	1039+1059.c	5B	140	Trình tự nucleotit của PDISP/H3 Perth với vòng phân giải protein được xóa bỏ	52A
40	Cátxét 1059	5C	141	H3P1609(Prl-)#2.r	52B
41	Trình tự axit amin PDISP/HA cúm B/Brisbane/60/2008 (vòng phân giải protein được xóa bỏ)	5D	142	H3P1609(Prl-)#2.c	52C
42	Trình tự nucleotit H5 A/Indonesia/5/2005	1G	143	Trình tự axit amin của PDISP/H3 Perth với	52D

SEQ ID NO:	Mô tả	Hình vẽ	SEQ ID NO:	Mô tả	Hình vẽ
	(H5N1)			vòng phân giải protein được xóa bỏ	
43	trình tự nucleotit PDISP/HA cúm B/Brisbane/60/2008 (vòng phân giải protein được xóa bỏ)	5E	144	Trình tự nucleotit của PDISP/H3 Victoria với vòng phân giải protein được xóa bỏ	53A
44	Trình tự tự nhiên vị trí tách H5/Indo	19	145	H3V36111(Prl-).r	53B
45	Vị trí tách cải biến H5/Indo (TETR)	19	146	H3V36111(Prl-).c	53C
46	Vị trí tách cải biến H5/Indo (TETQ)	19	147	Trình tự axit amin của PDISP/H3 Victoria với vòng phân giải protein được xóa bỏ	53D
47	Vị trí tách H1	Bảng 1	148	Trình tự nucleotit của PDISP/H7 Hangzhou	54A
48	Vị trí tách H3	Bảng 1	149	IF*-H7H113.s1-6r	54B
49	IF-HAB110.S1+3c	6A	150	Trình tự axit amin của PDISP/H7 Hangzhou	54C
50	IF-HAB110.s1-4r	6B	151	Trình tự nucleotit của PDISP/H7 Hangzhou với vòng phân giải protein được xóa bỏ	55A
51	HA B Wisconsin tổng hợp	6C	152	H7H113(PrL-).r	55B
52	Cấu trúc 193	6E	153	H7H113(PrL-).c	55C
53	Cátxét 1462	6F	154	Trình tự axit amin của PDISP/H7 Hangzhou với vòng phân giải protein được xóa bỏ	55D
54	Trình tự axit amin HA cúm B/Wisconsin/1/2010	6G	155	Trình tự nucleotit của PDISP/H9 Hong Kong	56A
55	HAB110(PrL-).r	7A	156	IF**-H9HK107399.S1-6r	56B
56	HAB110(PrL-).c	7B	157	Trình tự axit amin của PDISP/H9 Hong Kong	56C
57	Cátxét 1467	7C	158	Trình tự nucleotit của PDISP/H9 Hong Kong với vòng phân giải	57A

SEQ ID NO:	Mô tả	Hình vẽ	SEQ ID NO:	Mô tả	Hình vẽ
				protein được xóa bỏ	
58	Trình tự axit amin HA cúm B/Wisconsin/1/2010 (loại bớt PL)	7D	159	H9HK107399(Prl-).r	57B
59	Vị trí tách B	Bảng 1	160	H9HK107399(Prl-).c	57C
60	Vị trí tách tự nhiên H5/Indo	Bảng 2	161	Trình tự axit amin của PDISP/H9 Hong Kong với vòng phân giải protein được xóa bỏ	57D
61	Vị trí tách cải biến H5/Indo	Bảng 2	162	Trình tự nucleotit của PDISP/HA B Malaysia	58A
62	Vị trí tách cải biến H5/Indo	Bảng 2	163	IF**-HBM250604.S1-6r	58B
63	Vị trí tách cải biến H1/Brisbane	Bảng 2	164	Trình tự axit amin của PDISP/HA B Malaysia	58C
64	Vị trí tách cải biến H3/Brisbane	Bảng 2	165	Trình tự nucleotit của PDISP/HA B Malaysia với vòng phân giải protein được xóa bỏ	59A
65	Vị trí tách cải biến B/Florida, B/Brisbane	Bảng 2	166	HBM250604(PrL-).r	59B
66	A/H3/HA0 liên ứng		167	HBM250604(PrL-).c	59C
67	A/H1/HA0 liên ứng		168	Trình tự axit amin của PDISP/HA B Malaysia với vòng phân giải protein được xóa bỏ	59D
68	B/HA0 liên ứng		169	Trình tự nucleotit của PDISP/HA B Massachusetts	60A
69	xóa bỏ vòng phân giải protein H5 Anhui	18C	170	Trình tự axit amin của PDISP/HA B Massachusetts	60B
70	xóa bỏ vòng phân giải protein H5 Indo	18C	171	trình tự nucleotit của PatA SP	
71	xóa bỏ vòng phân giải protein H5 Vietnam	18C	172	trình tự axit amin của patatin A peptit tín hiệu	
72	xóa bỏ vòng phân giải protein B Florida	18C	173	Trình tự CPMV HT	
73	xóa bỏ vòng phân giải protein B Malaysia	18C	174	Trình tự kozak liên ứng - giới thực vật	

SEQ ID NO:	Mô tả	Hình vẽ	SEQ ID NO:	Mô tả	Hình vẽ
74	MutCleavage-H5(Indo).r	23A	175	Trình tự kozak liên ứng-thực vật hai lá mầm	
75	MutCleavage-H5(Indo).c	23B	176	Trình tự kozak liên ứng - arabidopsis	
76	Cátxét 676	23C	177	Trình tự kozak liên ứng	
77	Trình tự axit amin cúm A/Indonesia/5/2005 (H5N1) TETR đột biến vị trí tách.	23D	178	Trình tự kozak liên ứng	
78	H5I505_TETQ.r	24A	179	Trình tự kozak liên ứng	
79	H5I505_TETQ.c	24B	180	Trình tự kozak liên ứng	
80	Cátxét 766	25C	181	Trình tự kozak liên ứng	
81	Trình tự axit amin cúm A/Indonesia/5/2005 (H5N1) TETQ đột biến vị trí tách.	25D	182	Trình tự kozak liên ứng	
82	H5I505(PrL-).r	26A	183	Trình tự kozak liên ứng	
83	H5I505(PrL-).c	26B	184	Trình tự kozak liên ứng	
84	Cátxét 928	26C	185	Trình tự kozak liên ứng	
85	Trình tự axit amin cúm A/Indonesia/5/2005 (H5N1) với vòng phân giải protein được xóa bỏ.	26D	186	Trình tự liên ứng Kozak	
86	IF-S2+S4-B Bris.c	30A	187	Trình tự nucleotit của CPMV HT+	
87	IF-S1a4-B Bris.r	30B	188	Trình tự nucleotit của CPMV HT+ 511	
88	Gen HA B Brisbane tổng hợp	30C	189	Trình tự nucleotit của CPMV HT+[WT115]	
89	Cấu trúc 1029	30D	190	HBF406(PrL-).r	61A
90	Trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008	30E	191	HBF406(PrL-).c	61B
91	Trình tự nucleotit của PDISP/HA B Brisbane (PrL-).	31A	192	IF*-HBF406.s1-6r	61C
92	Trình tự axit amin của PDISP/HA B Brisbane (PrL-)	31B	193	Trình tự nucleotit của PDISP/HA B Florida với vòng phân giải protein được xóa bỏ	61D

SEQ ID NO:	Mô tả	Hình vẽ	SEQ ID NO:	Mô tả	Hình vẽ
93	Trình tự nucleotit của CPMVX/CPMVX+		194	Trình tự axit amin của PDISP/HA B Florida với vòng phân giải protein được xóa bỏ	61E
94	trình tự axit amin liên ứng của miền xuyên màng		195	Cátxét biểu hiện số 2102	61F
95	Trình tự nucleotit của PDISP/HA B Brisbane (PrL-)+H1 California TMCT.	33A	196	Cátxét biểu hiện số 2104	61H
96	Trình tự axit amin của PDISP/HA B Brisbane (PrL-)+H1 California TMCT.	33B	197	IF-H1cTMCT.S1-4r	62A
97	Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-)	34A	198	Trình tự nucleotit của PDISP/HA B Florida+H1Cal TMCT với vòng phân giải protein được xóa bỏ	62B
98	Trình tự axit amin của PDISP/HA B Massachussetts (PrL-)	34B	199	Trình tự axit amin của PDISP/HA B Florida+H1Cal TMCT với vòng phân giải protein được xóa bỏ	62C
99	Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT.	35A	200	Cátxét biểu hiện số 2106	62D
100	Trình tự axit amin của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT.	35B	201	Cátxét biểu hiện số 2108	62F
101	Trình tự nucleotit của HA B Wisconsin (PrL-).	36A			

Ví dụ thực hiện sáng chế

Ví dụ 1

Gây nhiễm agrobacterium

Chủng *Agrobacterium* AGL1 được gây nhiễm bằng chuyển gien bằng xung điện với cấu trúc ADN bằng cách sử dụng phương pháp được mô tả bởi D'Aoust et al 2008 (Plant Biotechnology Journal 6:930-940). *Agrobacterium* được gây nhiễm được sinh trưởng trong môi trường YEB được bổ sung với 10mM axit 2-(N-morpholino)etansulfonic (MES), 20µM axetosyringon, 50µg/ml kanamyxin và 25µg/ml carbenixilin độ pH=5,6 đến OD_{600} khoảng từ 0,6 đến 1,6. Huyền phù *Agrobacterium* được ly tâm trước khi sử dụng và tạo huyền phù lại trong môi trường thấm lọc (10mM $MgCl_2$ và 10mM MES độ pH=5,6).

Tạo ra sinh khối thực vật, tiêm truyền và thấm lọc nông nghiệp

Thuật ngữ "sinh khối" và "vật liệu thực vật" như được sử dụng trong bản mô tả này có nghĩa là phản ánh vật liệu bất kỳ thu được từ thực vật. Sinh khối hoặc vật liệu thực vật có thể bao gồm toàn bộ thực vật, mô, tế bào hoặc phân đoạn bất kỳ của nó. Hơn nữa, sinh khối hoặc vật liệu thực vật có thể bao gồm thành phần thực vật nội bào, thành phần thực vật ngoại bào, phần chất dạng lỏng hoặc dạng rắn của thực vật hoặc tổ hợp của chúng. Hơn nữa, sinh khối hoặc vật liệu thực vật có thể bao gồm thực vật, tế bao thực vật, mô, phần chiết dạng lỏng hoặc tổ hợp của chúng, từ lá thực vật, thân, quả, rễ hoặc tổ hợp của chúng. Phần của thực vật có thể bao gồm vật liệu thực vật hoặc sinh khối.

Thực vật *Nicotiana benthamiana* được sinh trưởng từ hạt trong nền phẳng được làm đầy với chất than rêu có bán trên thị trường. Thực vật được cho phép sinh trưởng trong nhà kính trong chu kỳ sáng 16/8 và chế độ nhiệt độ 25°C ngày/20°C đêm. Ba tuần sau khi nảy mầm, cây non riêng rẽ được chọn, được ghép vào chậu và cho phép sinh trưởng trong nhà kính trong ba tuần nữa trong cùng điều kiện môi trường.

Agrobacteria được gây nhiễm với mỗi cấu trúc được sinh trưởng trong môi trường YEB được bổ sung với 10mM axit 2-(N-morpholino)etansulfonic (MES), 20µM axetosyringon, 50µg/ml kanamycin và 25µg/ml carbenixilin độ pH=5,6 cho đến khi chúng

đạt đến OD_{600} khoảng từ 0,6 đến 1,6. Huyền phù *Agrobacterium* được ly tâm trước khi sử dụng và tạo huyền phù lại trong môi trường thấm lọc (10mM $MgCl_2$ và 10mM MES độ pH=5,6) và lưu trữ qua đêm ở 4°C. Vào ngày thấm lọc, mẻ nuôi cấy được pha loãng với thể tích nuôi cấy 2,5 và cho phép làm ấm trước khi sử dụng. Thực vật nguyên vẹn của *N. benthamiana* được đặt lộn ngược vào huyền phù vi khuẩn trong bể bằng thép không gỉ kín khí trong chân không từ 20 đến 40 Torr trong 2 phút. Thực vật được cho trở lại nhà kính trong khoảng thời gian từ 2 đến 6 ngày ủ cho đến khi thu hoạch.

Thu hoạch lá và tổng phần chiết protein

Sau khi ủ, phần trên không khí của thực vật được thu hoạc, làm đông lạnh ở -80°C và nghiền thành các mẩu nhỏ. Tổng protein có thể hoà tan được chiết bằng cách làm đồng nhất (Polytron) mỗi mẫu vật liệu thực vật được nghiền được làm đông lạnh trong 3 thể tích 50mM Tris mát độ pH=8,0, 0,5M NaCl, 0,1% Triton X-100 và 1mM phenylmetansulfonyl florua. Sau khi làm đồng nhất, ly tâm bùn nhão ở 10.000g trong 10 phút ở 4°C và phần chiết thô sạch này (dịch nổi) được giữ để phân tích.

Phân tích protein và thấm tách miễn dịch

Hàm lượng protein của phần chiết thô sạch được xác định bằng thử nghiệm Bradford (Bio-Rad, Hercules, CA) bằng cách sử dụng albumin huyết thanh bò làm tiêu chuẩn tham chiếu. Protein được tách bằng SDS-PAGE và truyền điện vào màng polyvinylen diflorua (PVDF) (Roche Diagnostics Corporation, Indianapolis, IN) để phát hiện miễn dịch. Trước khi thấm tách miễn dịch, màng được phong bế bằng 5% sữa không kem và 0,1% Tween-20 trong nước muối đệm Tris (TBS-T) trong từ 16 đến 18 giờ ở 4°C.

Việc thấm tách miễn dịch được thực hiện với việc ủ lần thứ nhất bằng kháng thể sơ cấp (Bảng 4 thể hiện kháng thể và điều kiện được sử dụng để phát hiện mỗi HA), trong 2µg/ml trong 2% sữa không kem trong TBS-Tween 20 0,1%. Kháng thể thứ cấp được sử dụng để phát hiện phát quang hoá học được chỉ ra trong Bảng 4, được pha loãng như được chỉ định theo 2% sữa không kem trong TBS-Tween 20 0,1%. Phức hệ phản ứng miễn dịch được phát hiện bằng phát quang hoá học bằng cách sử dụng luminol làm chất nền (Roche Diagnostics Corporation). Thể tiếp hợp horseradish peroxidaza–enzym của kháng thể IgG

của người đượ thực hiện bằng cách sử dụng kit tiếp hợp EZ-Link Plus® Activated Peroxidase (Pierce, Rockford, IL).

Bảng 4: Các điều kiện điện di, kháng thể và sự pha loãng của thấm tách miễn dịch của protein được biểu hiện.

Typ phụ HA	Chủng cúm	Điều kiện điện di	Kháng thể sơ cấp	Pha loãng	Kháng thể thứ cấp	Pha loãng
B	B/Brisbane/60/2008	Không khử	TGA, AS397	1:20000	Kháng thể thỏ kháng cừu (JIR 313-035-045)	1:10 000
B	B/Wisconsin/1/2010	Không khử	NIBSC 07/356	1 :2000	Kháng thể thỏ kháng cừu (JIR 313-035-045)	1:10 000
B	B/Malaysia/2506/2004	Không khử	NIBSC 07/184	1 :2000	Kháng thể thỏ kháng cừu (JIR 313-035-045)	1:10 000
H3	A/Perth/16/2009 (H3N2)	Không khử	TGA, AS400	1:20000	Kháng thể thỏ kháng cừu (JIR 313-035-045)	1:10 000
H3	A/Victoria/361/2011	Không khử	TGA, AS400	1 :20000	Kháng thể thỏ kháng cừu (JIR 313-035-045)	1:10 000
H1	A/California/07/2009 (H1N1)	Khử	Sino, 11055-MMO1	1 µg/ml	Dê kháng chuột (JIR 115-035-146)	1:7 500

H5	A/Indonesia/05/2005 (H5N1)	Khử	CBER, S-7858	1:4000	Kháng thể thỏ kháng cừu (JIR 313-035-045)	1:10 000

JIR: Jackson ImmunoResearch, West Grove, PA, USA;

CBER: Center for Biologics Evaluation và Research, Rockville, MD, USA.

Sino: Sino Biological inc., Beijing, China.

TGA: Therapeutic Goods Administration, Australia.

NIBSC: National Institute for Biological Standards and Control, United Kingdom

Thử nghiệm hemagglutin hoá

Thử nghiệm hemagglutin hoá được dựa vào phương pháp được mô tả bởi Nayak and Reichl (2004). Tóm lại, dãy pha loãng hai lần của mẫu thử nghiệm (100µL) được tạo thành trong đĩa vi chuẩn 96 giếng đáy hình chữ V chứa 100µL PBS, cho phép 100µL mẫu pha loãng/giếng. 100 microlit của 0,25% huyền phù tế bào máu đỏ của gà tây (Bio Link Inc., Syracuse, NY) được bổ sung vào mỗi giếng và ủ đĩa trong 2 giờ ở nhiệt độ trong phòng. Sự tương hỗ của phần pha loãng lớn nhất thể hiện sự hemagglutin hoá hoàn thiện được ghi chép như hoạt tính HA. Theo cách song song, tiêu chuẩn HA tái tổ hợp (A/Vietnam/1203/2004 H5N1) (Protein Science Corporation, Meriden, CT) được pha loãng trong PBS và chạy như đối chứng trên mỗi đĩa.

Phần chiết VLP bằng việc tiêu hoá vách tế bào

Mô lá được gom từ thực vật *Nicotiana benthamiana* và cắt thành các mẩu từ $1cm^2$. Các mẩu lá được ngâm trong 500mM mannitol trong 30 phút ở nhiệt độ trong phòng (RT). Sau đó, loại bỏ dung dịch manitol và thay đổi với hỗn hợp enzym (hỗn hợp gồm xelulaza từ *Trichoderma viride* (Onozuka R-10; 3% thể tích) và hỗn hợp gồm pectinaza từ *Rhizopus* sp. (MACEROZYME™; 0,75% thể tích; cả hai của Yakult Pharmaceuticals) trong dung dịch nguyên sinh (500mM manitol, 10mM $CaCl_2$ và 5mM MES/KOH (độ pH=5,6)). Tỷ lệ được sử dụng bằng 20g mẩu lá/100mL dung dịch. Chế phẩm này được lan

toả theo cách đều nhau vào trong bình cạn (khoảng 11x18cm) và ủ trong 16 giờ trên bộ lắc quay ở 40 vòng/phút và 26°C.

Theo cách khác, phần chiết VLP có thể được thực hiện như sau: thực vật được thấm lọc nông nghiệp bằng AGL1/#489, 928, 676 và 766. Mô lá được gom từ thực vật *N. benthamiana* ở ngày 7 sau thấm lọc và cắt thành các mẩu nhỏ khoảng 1cm^2. Pectinaza 162L (Biocatalysts), Multifect CX CG và Multifect CX B (Genencor) được bổ sung vào 200mM Manitol, 75mM xitrat, 0,04% chất đệm natri bisulfua độ pH=6,0; chất đệm tiêu hoá. Sinh khối được tiêu hoá hai lần qua đêm ở nhiệt độ trong phòng trong thiết bị lắc quãy đạo.

Sau khi chiết trợ giúp enzym, loại bỏ các mẩu lá bằng việc lọc (bộ lọc ni lông mắt lưới 250 hoặc 400µm). Ly tâm phần chiết lọc thô ở 5000xg trong 5 phút. Dịch nổi được cho vào phát hiện sự biểu hiện HA (hoạt tính hemagglutin hoá (xem Fig.20) và thấm tách Western (xem Fig.22).

Ví dụ 2:

Tác dụng của vòng phân giải protein được cải biến đối với sự kết tụ của HA

Như được thể hiện trên Fig.13A, sự biểu hiện của B/Brisbane (cấu trúc số:1008) nguyên thể là thấp hơn so với sự biểu hiện của B/Brisbane bao gồm vòng phân giải protein được cải biến (cấu trúc số 1059). Hoạt tính hemagglutin hoá gia tăng cũng được quan sát với B/Brisbane bao gồm vòng phân giải protein được cải biến (cấu trúc số 1059) khi so với B/Brisbane HA nguyên thể (cấu trúc số 1008; Fig.13B).

Các kết quả tương tự được quan sát với mức kết tụ của B/ Wisconsin bao gồm vòng phân giải protein được cải biến (cấu trúc số 1467), mà lớn hơn so với mức kết tụ được quan sát đối với B/Wisconsin HA nguyên thể (cấu trúc số 1462; Fig.16A). Hoạt tính hemagglutin hoá gia tăng cũng được quan sát với B/Wisconsin bao gồm vòng phân giải protein được cải biến (cấu trúc số 1467) khi so với B/Wisconsin HA nguyên thể (cấu trúc số 1462; Fig.16B) cho biết sự kết tụ lớn hơn đối với protein đột biến.

Sự biểu hiện của H5/Indo bao gồm vòng phân giải protein được cải biến cũng được quan sát với các cải biến bao gồm vòng phân giải protein bao gồm cầu liên kết GG (cấu trúc số 928; SEQ ID NO:85), cầu liên kết TETR (cấu trúc số 676; SEQ ID NO:77), hoặc cầu liên kết TETQ (cấu trúc số 766; SEQ ID NO: 8; Fig.22).

Tác dụng của sự đồng biểu hiện M2 cúm đối với mức kết tụ HA

Sự đồng biểu hiện của M2 được đánh giá đối với ảnh hưởng của nó đối với mức kết tụ của cúm B HA được cải biến. Cấu trúc số 1059 mã hoá cúm B HA trong đó vòng phân giải protein được thay thế bằng 2 cầu liên kết axit amin (GG thay thế cho aa 341-359; SEQ ID NO: 41). Các kết quả từ phép phân tích thấm tách western được thể hiện trên Fig.13A thể hiện rằng việc loại bỏ vòng phân giải protein dẫn đến mức kết tụ cúm B HA gia tăng (so sánh 1008 với 1059) và mà sự đồng biểu hiện của M2 với cúm B HA cải biến cúng gia tăng mức kết tụ HA (Fig.13A, 1059 đối lại 1059+1261). Phép phân tích hoạt tính hemagglutin hoá đối với phần chiết protein thô từ thực vật được biến nạp với cúm B HA có hoặc không có sự cải biến và có hoặc không có sự đồng biểu hiện của M2 xác nhận tác dụng tích cực của sự đồng biểu hiện M2 đối với mức kết tụ của cúm B HA nguyên thể (Fig.13B, 1008 đối lại 1008+1261) và cúm B HA cải biến (Fig.13B, 1059 đối lại 1059+1261).

Sự đồng biểu hiện của M2 với typ A HA bao gồm vòng phân giải protein được cải biến cũng dẫn đến sự biểu hiện của HA. Ví dụ, sự đồng biểu hiện của H3 cải biến, với vòng phân giải protein được thay thế bằng cầu liên kết GS hoặc cầu liên kết $(GSS)_3$ (xem Fig.21E, SEQ ID NO: 20; Fig.21F, SEQ ID NO: 21), cùng với M2 cũng có thể dẫn đến sự kết tụ HA ở thực vật.

Hiệu quả của M2 từ cúm A/Puerto Rico/8/1934 làm gia tăng sự kết tụ của cúm B HA và H3 cải biến được so với sự kết tụ của M2 từ cúm A/New Caledonia/20/1999. Đối với cúm B HA cải biến, phép so sánh được thực hiện bằng phép phân tích thấm tách western của phần chiết protein từ thực vật được biến nạp với các cấu trúc 1059, 1059+1261 và 1059+859. Các kết quả thu được chứng minh rằng sự đồng biểu hiện của M2 từ cúm A/Puerto Rico/8/1934 (được mã hoá bởi cấu trúc số 859) là hữu hiệu như sự

đồng biểu hiện của M2 từ cúm A/New Caledonia/20/1999 (được mã hoá bởi cấu trúc số 1261) để gia tăng sự kết tụ của cúm B HA cải biến (Fig.14).

Tác dụng của sự đồng biểu hiện cúm M2 đối với mức kết tụ của các chủng B HA khác nhau

Phép phân tích thấm tách Western của phần chiết protein từ thực vật được biến nạp với các cấu trúc gen điều khiển sự biểu hiện của cúm B HA (từ B/Wisconsin/1/2010) (cấu trúc số 1462) với sự có mặt hoặc không có mặt của cấu trúc biểu hiện M2 (cấu trúc số 1261) thể hiện rằng sự đồng biểu hiện M2 dẫn đến sự kết tụ gia tăng của cúm B HA (Fig.16A).

Sự đồng biểu hiện của M2 cũng được đánh giá đối với sự tác động của nó đối với mức kết tụ của cúm B HA cải biến. Cấu trúc số 1467 mã hoá cúm B HA trong đó vòng phân giải protein được thay thế bằng cầu liên kết 2 axit amin (GG thay thế cho aa 341-359; SEQ ID NO: 58). Các kết quả từ phép phân tích thấm tách western được thể hiện trên các hình vẽ Fig.16A thể hiện việc loại bỏ vòng phân giải protein dẫn đến mức kết tụ cúm B HA gia tăng (so sánh 1462 với 1467) và mà sự đồng biểu hiện của M2 với cúm B HA cải biến cũng làm tăng mức kết tụ HA (Fig.16A, 1467 đối lại 1467+1261). Phép phân tích của hoạt tính hemagglutin hoá đối với phần chiết protein thô từ thực vật được biến nạp với cúm B HA có hoặc không có sự cải biến và có hoặc không có sự đồng biểu hiện của M2 xác nhận tác dụng tích cực của sự đồng biểu hiện M2 đối với mức kết tụ của cúm B HA nguyên thể (Fig.16B, 1462 đối lại 1462+1261) và cúm B HA cải biến (Fig.16B, 1467 đối lại 1467+1261).

Tác dụng của yếu tố khuếch đại BeYDV và vòng phân giải protein được cải biến đối với sự kết tụ của HA

Phép phân tích thấm tách Western của phần chiết protein từ thực vật được biến nạp với cấu trúc gen điều khiển sự biểu hiện của cúm B HA cải biến (từ B/Brisbane/60/2008) có hoặc không có vòng phân giải protein được loại bỏ (xem Fig.17A đối với cấu trúc) và với sự có mặt hoặc không có mặt của yếu tố khuếch đại BeYDV (cấu trúc số 1059 và

1039) thể hiện rằng với sự không có mặt của BeYDV không có sự kết tụ của cúm B HA có thể được phát hiện (Fig.17B), khi yếu tố điều hoà là CPMV-HT.

Tác dụng của vòng phân giải protein được cải biến đối với độ chuẩn HA tương đối và hemagglutin hoá

Với việc tham khảo đến Fig.29A, đã thể hiện phép so sánh của hoạt tính của protein HA cải biến được tạo ra ở thực vật bao gồm CPMV HT, CPMV HT+, CPMV 160 hoặc CPMV160+, dựa vào yếu tố tăng cường được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá HA cải biến với vòng phân giải protein được phát hiện (được thay thế bằng cầu liên kết GG) hoặc HA nguyên thể. Trong hầu hết các trường hợp, sự biểu hiện (được xác định như độ chuẩn hemagglutin hoá hoặc hoạt tính) là cao hơn đối với các cấu trúc được tạo thành chủ yếu từ CPMV HT+, CPMV 160 hoặc CPMV160+ chứng minh mức biểu hiện đáng kể.

Bảng 5a: Độ chuẩn HA tương đối (HA=1 kiểu hoang dại) (xem Fig.29A)

	Fct = 1					
	Kiểu hoang dại			PrL-		
Mô tả (n=)	GM	MD	n=	GM	MD	n=
H2 Sin157 (HT+) (n = 2)	1,0	0,4	2	1,2	0,1	2
H2 Sin157 (160+) (n = 2)	1,0	0,3	2	1,9	0,9	2
H3 Per1609 (HT+) (n = 4)	1,0	0,0	4	4,9	4,2	4
H3 Per1609 (160+) (n = 4)	1,0	0,3	4	6,1	3,6	4
H3 Vic36111 (HT+) (n = 6)	1,0	0,3	6	0,9	0,3	6
H3 Vic36111 (160+) (n = 6)	1,0	0,3	6	1,2	0,4	6
H5 Indo505 (HT) (n = 3)	1,0	0,1	3	1,2	0,3	3
H7 Han113 (HT+) (n = 5)	1,0	0,2	5	1,0	0,1	5
H9 HK107399 (HT+) (n = 2)	1,0	0,4	2	1,2	0,2	2
H9 HK107399 (160+) (n = 2)	1,0	0,2	2	0,7	0,1	2
HB Bri6008 (HT) (n = 1)	1,0	0,6	1	0,6	0,1	1
HB Mal250604 (160+) (n = 2)	1,0	0,3	2	8,4	2,4	2
HB Mas212 (HT) (n = 3)	1,0	0,1	3	1,4	1,3	3
HB Mas212 (HT+) (n = 3)	1,0	0,6	3	6,7	3,5	3
HB Mas212 (160+) (n = 2)	1,0	0,8	2	11,5	4,2	2

Ví dụ 3

Sản lượng H7 Hangzhou HA VLP gia tăng khi vòng phân giải protein được loại bỏ (PrL-) so với cấu trúc nguyên thể.

Thực vật *N. benthamiana* được thấm lọc với AGL1/#2142+1261 và #2152+1261 và lá được thu hoạch sau thời gian ủ bảy ngày. Gom mô lá và cắt thành các mẩu khoảng 1cm^2. Pectinase 162L và Pectinase 444L (Biocatalysts), Multifect CX CG và Multifect CX B (Genencor) được bổ sung vào với 200mM Manitol, 125mM xitrat, 0,04% chất đệm natri bisulfit độ pH=6,0. Tiêu hoá sinh khối qua đêm ở nhiệt độ trong phòng trong thiết bị lắc quỹ đạo.

Sau khi tiêu hoá, lọc phân đoạn không hợp bào qua bộ lọc 400µm ni lông để loại bỏ mô sinh dưỡng thô không tiêu hoá (<5% sinh khối khởi đầu). Sau đó, ly tâm phần chiết lọc ở nhiệt độ trong phòng trong 15 phút ở 5000xg để loại bỏ thể nguyên sinh và chất gây nhiễm nội bào (protein, ADN, màng, túi, sắc tố, v.v). Tiếp theo, dịch nổi được lọc kỹ (để làm sạch) bằng cách sử dụng sợi thuỷ tinh 1,2µm (Sartopore GF plus/Sartorius Stedim), và bộ lọc 0,45/0,2µm (Sartopore 2/Sartorius Stedim), trước khi được cho vào sắc ký.

Nạp phân đoạn không hợp bào sạch vào cột trao đổi cation (Poros HS Applied Biosystems) được làm cân bằng với chất đệm cân bằng/tách rửa (50mM NaPO4, 100mM NaCl, 0,005% Tween 80 độ pH=6,0). Khi tia cực tím trở lại zero, bước tách rửa phần chiết bằng chất đệm cân bằng/tách rửa chứa nồng độ gia tăng NaCl (500mM). VLP tinh chế được cô đặc bằng TFF, lọc pha loãng đối lại chất đệm bào chế (100mM PO4, 150mM NaCl, 0,01% Tween 80 ở độ pH=7,4) và chuyển qua bộ lọc 0,22µm.

Thử nghiệm hemagglutin hoá đối với H7 được thực hiện dựa trên phương pháp được mô tả bởi Nayak và Reichl (2004). Tóm lại, phần pha loãng hai lần liên tiếp của mẫu thử nghiệm (100µL) được thực hiện trong đĩa vi chuẩn 96 giếng đáy hình chữ V chứa 100µL PBS, còn lại 100µL mẫu pha loãng/giếng. 100 microlit của 0,25% huyền phù tế bào hồng huyết cầu gà tây (Bio Link Inc., Syracuse, NY) được bổ sung vào mỗi giếng và ủ đĩa trong 2 giờ ở nhiệt độ trong phòng. Sự thuận nghịch của phần pha loãng cao nhất thể hiện sự hemagglutin hoá hoàn toàn được ghi chép như hoạt tính hemagglutin hoá.

Tổng hàm lượng protein của phần chiết thô được gạn được xác định bằng cách sử dụng albumin huyết thanh bò làm tiêu chuẩn tham khảo. Sản lượng tương đối thu được bằng cách so sánh PrL-cấu trúc với cấu trúc nguyên thể được sử dụng làm đối chứng. Việc

tách bằng SDS-PAGE, với việc làm biến tính mẫu nạp chất đệm (0,1M Tris độ pH=6,8, 0,05% xanh bromophenol, 12,5% glyxerol, 4% SDS và 5% beta-mercaptoetanol), được thực hiện trong các điều kiện khử và Coomassie Brillant Blue R-250 được sử dụng để nhuộm protein.

Fig.46A thể hiện rằng hoạt tính hemagglutin hoá trong phần chiết thực vật là lớn hơn đối với cấu trúc H7 Hangzhou trong đó vòng phân giải protein được loại bỏ (#2152 + #1261, xem ví dụ 5.34) so với cấu trúc nguyên thể (# 2142 + # 1261, xem ví dụ 5.33).

Fig.46B thể hiện mà tổng sản lượng protein tương đối trong VLP được tinh chế là lớn hơn đối với cấu trúc H7 Hangzhou trong đó vòng phân giải protein được loại bỏ (#2152 + #1261) so với cấu trúc nguyên thể (# 2142 + # 1261). Ví dụ này chứng minh sự tương quan tốt giữa sự cải thiện VLP được kết tụ ở thực vật đối lại sản lượng cuối cùng khi thực hiện quy trình hoàn thiện.

Fig.46C thể hiện phép phân tích SDS-PAGE, với dãy 2 thể hiện cấu trúc H7 Hangzhou được tinh chế với vòng phân giải protein được loại bỏ và dãy 3 thể hiện cấu trúc H7 Hangzhou nguyên thể được tinh chế. Đối với mỗi dãy, 2µg tổng protein được nạp vào gel. Độ tinh khiết của profin protein là tương tự đối với cả hai cấu trúc và lớn hơn 90%.

Ví dụ 4.1.
Sức kháng trypsin của thể đột biến H5 Indonesia VLP trong đó vòng phân giải protein được cải biến hoặc được loại bỏ là lớn hơn so với H5 Indonesia nguyên thể.

Thực vật *N. benthamiana* được thấm lọc nông nghiệp với AGL1/#489, #928, #766 và # 676 như được mô tả trong Ví dụ 1 (trên đây). Gom lá từ các thực vật 7 ngày sau thấm lọc, cắt thành các mẩu khoảng 1cm^2. Pectinase 162L (Biocatalysts), Multifect CX CG và Multifect CX B (Genencor) được bổ sung vào với 200mM Manitol, 75mM xitrat, 0,04% chất đệm natri bisulfit độ pH=6,0. Sinh khối được tiêu hoá qua đêm ở nhiệt độ trong phòng trong thiết bị lắc quỹ đạo. Lọc thô phần chiết tiêu hoá, ly tâm, gạn và tinh chế như được mô tả trong Ví dụ 3 (H7 Hangzhou).

Đối với mỗi (# 489), PRL- (# 928), TETQ (# 766) và TETR (# 676) nguyên thể, phần chiết H5 Indonesia HA VLP, hai mẫu HA VLP được tạo huyền phù lại trong chất đệm (100mM Na/KPO_4, 150mM NaCl, 0,01% TWEEN 80) ở độ pH=7,4. Trypsin được bổ sung vào với tỷ lệ 1:100 protein. Thu được các mẫu sau 30, 60 và 120 phút ủ ở nhiệt độ trong phòng, sau đó đun sôi trong chất đệm nạp mẫu để làm ngừng phản ứng. Phần chiết không tiêu hoá (đối chứng) và phần chiết tiêu hoá trypsin được phân tích bằng SDS-PAGE gel như được mô tả trong Ví dụ 3.

Fig.47A thể hiện phép phân tích SDS-PAGE của mẫu tiêu hoá trypsin, với các dãy từ 2 đến 5 thể hiện H5 Indonesia VLP (#489) nguyên thể, với các dãy từ 6 đến 9 thể hiện PrL- H5 Indonesia VLP (#928), với các dãy từ 10 đến 13 thể hiện TETQ H5 Indonesia VLP (#766) và với các dãy từ 14 đến 17 thể hiện TETR H5 Indonesia VLP (#676) ở các thời điểm khác nhau trong tiêu hoá (0, 30, 60, và 120 phút). H5 Indonesia VPL nguyên thể, với dải tương ứng với HA0 monome là có thể phát hiện ở khoảng 75kDa trong phần chiết không tiêu hoá trong dãy 2, được xử lý một cách nhanh chóng thành các dải HA1 và HA2 qua việc bổ sung trypsin, có thể phát hiện ở khoảng 50 và 25kDa một cách tương ứng trong quá trình tiêu hoá trypsin trong các dãy từ 3 đến 5. PrL- và TETQ H5 Indonesia VLP, được làm ổn định bằng việc loại bỏ hoặc cải biến vị trí phân giải protein, thể hiện sức kháng trypsin như dải HA0 không tách thành các dải HA1 và HA2. TETR H5 Indonesia VLP ban đầu được làm ổn định bằng việc cải biến vị trí phân giải protein và HA0 monome được tách thành HA1 và HA2 thấp hơn ở H5 Indonesia VLPS nguyên thể.

Các dữ liệu này chứng minh rằng việc bảo vệ thành công của HA0 protein ở vị trí phân giải protein của nó nằm trong HA1-HA2, bằng cách xóa bỏ vòng phân giải protein (prl-) hoặc thay thế vòng phân giải protein bằng phương pháp trình tự liên kết (TETQ).

Ví dụ 4.2. Tính gây miễn dịch của H5 Indonesia VLP nguyên gốc là tương tự với bản sao đột biến của nó (PrL-, TETQ và TETR) ở chuột.

Phần chiết PrL-, TETR và TETQ H5 Indonesia VLP nguyên thể được tinh chế như được mô tả trong Ví dụ 4.1 (trên đây).

Fig.47B thể hiện tính gây miễn dịch (độ chuẩn HI) của H5 VLP nguyên thể và bản sao thể đột biến (prl-, TETQ và TETR) ở chuột sau hai liều. Chuột BALB/c (n=8/nhóm) được tiêm hai lần trong cơ, 21 ngày sau, với 10ug liều vacxin H5 VLP được tạo thành chủ yếu từ thực vật (nguyên thể, prl-, TETQ hoặc TETR) dựa vào hàm lượng HA của nó. Phép phân tích độ chuẩn HI được thực hiện từ huyết thanh của mỗi động vật, 42 dpv (21 ngày sau liều thứ hai) và H5 VLP A/Indonesia/5/2005 (H5N1) được sử dụng làm kháng nguyên. Các cột thể hiện phép so sánh các độ chuẩn (%) HI tương đối của mỗi thể đột biến H5 VLP với H5 VLP nguyên thể (được tính toán với log2 của độ chuẩn HI GMT và 95% CI). Vi sai thống kê giữa các nhóm đối với mỗi liều được so sánh bằng cách sử dụng ANOVA một phương pháp tiếp theo bằng phép phân tích Tukey's post-hoc đối với các dữ liệu biến đổi Log2 (giả định sự phân bố thông thường của chúng). *$p < 0,05$ được xem là có nghĩa. Không có sự khác biệt giữa các nhóm đối với mỗi liều được quan sát.

Ví dụ 5.1: B-2X35S/CPMV-HT/M2 New Caledonia/NOS (Cấu trúc số 1261)

Trình tự mã hoá M2 từ cúm A/New Caledonia/20/1999 (H1N1) được tạo dòng vào trong hệ biểu hiện 2X35S/CPMV-HT/NOS ở plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR sau đây. Đoạn chứa trình tự mã hoá M2 hoàn thiện được khuếch đại bằng cách sử dụng các đoạn mồi IF-S1-M1+M2ANC.c (Fig.2A, SEQ ID NO: 7) và IF-S1-4-M2ANC.r (Fig.2B, SEQ ID NO: 8) bằng cách sử dụng gen M2 tổng hợp (tương ứng với nt 1-26 ghép cặp với nt 715-982 từ Số truy cập Genbank DQ508860; Fig.2C, SEQ ID NO: 9) làm mẫu. Sản phẩm PCR được tạo dòng vào hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1191 (Fig.1C) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1191 là plasmit nhận được dự định để tạo dòng "In Fusion" gen quan tâm trong cátxét biểu hiện dựa trên CPMV HT. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ bên trái sang bên phải các đường biên t-ADN được thể hiện trên Fig.1D (SEQ ID NO: 4). Cấu trúc thu được được ký hiệu là số 1261 (Fig.2D, SEQ ID NO: 10). Trình tự axit amin của M2 từ cúm

A/New Caledonia/20/1999 (H1N1) được thể hiện trên Fig.2E (SEQ ID NO: 11). Phép biểu diễn của plasmit 1261 được thể hiện trên Fig.11.

Ví dụ 5.2: C-2X35S/CPMV-HT/M2 Puerto Rico/NOS (Cấu trúc số 859)

Trình tự mã hoá M2 từ cúm A/Puerto Rico/8/1934 (H1N1) được tạo dòng vào trong hệ biểu hiện 2X35S/CPMV-HT/NOS trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR sau đây. Đoạn chứa trình tự mã hoá M2 hoàn thiện được khuếch đại bằng cách sử dụng các đoạn mồi IF-S1-M1+M2ANC.c (Fig.2A, SEQ ID NO: 7) và IF-S1-4-M2ANC.r (Fig.2B, SEQ ID NO: 8), bằng cách sử dụng gen M2 tổng hợp (tương ứng với nt 26-51 ghép nối với nt 740-1007 từ Số truy cập Genbank EF467824) (Fig.3A, SEQ ID NO: 12) làm mẫu. Sản phẩm PCR được tạo dòng trong hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1191 (Fig.1C) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1191 là plasmit nhận được dự định để tạo dòng "In Fusion" gen quan tâm trong cátxét biểu hiện dựa trên CPMV-HT. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Vật truyền là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang phía bên phải đường biên t-ADN được thể hiện trên Fig.1D (SEQ ID NO: 4). Cấu trúc thu được được ký hiệu số 859 (Fig.3B, SEQ ID NO: 13). Trình tự axit amin của M2 từ cúm A/Puerto Rico/8/1934 (H1N1) được thể hiện trên Fig.3C (SEQ ID NO: 14). Phép biểu diễn của plasmit 859 được thể hiện trên Fig.17.

Ví dụ 5.3: G-2X35S/CPMV-HT/PDISP/HA B Brisbane/NOS vào trong hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1008)

Việc tạo ra cấu trúc 1008 được mô tả trong patent Mỹ số US 61/541,780. Tóm lại, trình tự mã hoá HA từ cúm B/Brisbane/60/2008 được tạo dòng vào trong 2X35S/CPMV-HT/PDISP/NOS bao gồm hệ khuếch đại BeYDV+replicaza trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR bằng cách sử dụng gen HA B Brisbane tổng hợp (tương ứng với nt 34-1791 từ Số truy cập Genbank

FJ766840). Sản phẩm PCR được tạo dòng trong khung với peptit tín hiệu alfalfa PDI trong cátxét biểu hiện 2X35S/CPMV-HT/NOS vào trong hệ khuếch đại BeYDV. Cấu trúc 1194 (xem các hình vẽ Fig.4A, Fig.4B) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp để tạo ra cấu trúc số 1008 (Fig.4C, Fig. 9; SEQ ID NO: 32).

Cấu trúc số 1194 (Fig.4A) là plasmit nhận được dự định để tạo dòng "In Fusion" gen quan tâm trong khung với peptit tín hiệu alfalfa PDI trong cátxét biểu hiện dựa vào CPMV-HT vào trong hệ khuếch đại BeYDV. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA.

Ví dụ 5.4: I-2X35S/CPMV-HT/PDISP-HA B Brisbane với vòng phân giải protein được xóa bỏ vào trong hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1059)

Việc tạo ra cấu trúc 1059 được mô tả trong patent Mỹ số 61/541,780. Tóm lại, trình tự mã hoá HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ được tạo dòng vào trong 2X35S/CPMV-HT/PDISP/NOS bao gồm hệ khuếch đại BeYDV+replicaza trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp thắt dựa trên PCR (Darveau et al., 1995, Methods in Neuroscience 26: 77-85). Trong vòng thứ nhất của PCR, mảnh chứa trình tự mã hoá HA B Brisbane từ nt 46 đến nt 1065 được khuếch đại bằng cách sử dụng gen HA B Brisbane tổng hợp (tương ứng với nt 34-1791 từ số truy cập Genebank FJ766840) làm mẫu. Mảnh thứ hai, chứa trình tự mã hoá HA B Brisbane từ nt 1123 đến nt 1758, được khuếch đại bằng cách sử dụng gen HA B Brisbane tổng hợp (tương ứng với nt 34-1791 từ Số truy cập Genbank FJ766840) làm mẫu. Sau đó, sản phẩm PCR từ cả hai lần khuếch đại được trộn lẫn và sử dụng làm mẫu trong vòng khuếch đại thứ hai. Mảnh thu được (mã hoá HA B/Brisbane/60/2008 Δa.a. 356-374 với cầu liên kết GG giữa các mảnh; xem Fig.21B) được tạo dòng trong khung với peptit tín hiệu alfalfa PDI trong cátxét biểu hiện 2X35S/CPMV-HT/NOS bao gồm hệ khuếch đại BeYDV để tạo ra cấu trúc 1194 (Fig.4A, Fig.4B) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được

sử dụng cho phản ứng lắp ráp. Cấu trúc thu được được ký hiệu số 1059 (Fig.5C; SEQ ID NO: 40).

Trình tự axit amin của PDISP-HA B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ được thể hiện trên Fig.5D (SEQ ID NO: 41).

Ví dụ 5.5: B-2X35S/CPMV-HT/HA B Wisconsin/NOS vào trong hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1462)

Việc tạo ra cấu trúc 1462 được mô tả trong patent Mỹ số 61/541,780. Tóm lại, trình tự mã hoá HA từ cúm B/Wisconsin/1/2010 được tạo dòng vào trong 2X35S/CPMV-HT/NOS bao gồm hệ khuếch đại BeYDV(m)+replicaza trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR. Mảnh chứa trình tự mã hoá HA B Wisconsin hoàn thiện được tinh chế bằng cách sử dụng gen HA B Wisconsin tổng hợp (Số truy cập Genbank JN993010) làm mẫu. Sản phẩm PCR được tạo dòng trong cátxét biểu hiện 2X35S/CPMV-HT/NOS vào trong hệ khuếch đại BeYDV(m). Cấu trúc 193 (Fig.6D) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp.

Cấu trúc số 193 (Fig.6D, Fig.6E) là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong cátxét biểu hiện được tạo thành chủ yếu từ CPMV-HT vào trong hệ khuếch đại BeYDV(m). Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên Fig.6E (SEQ ID NO: 52). Cấu trúc thu được được ký hiệu số 1462 (Fig.6F, SEQ ID NO: 53). Trình tự axit amin của PDISP/HA từ cúm B/Wisconsin/1/2010 được thể hiện trên Fig.6G (SEQ ID NO: 54). Phép biểu diễn plasmit 1462 được thể hiện trên Fig.6H.

Ví dụ 5.6: C-2X35S/CPMV-HT/HA B Wisconsin với vòng phân giải protein được xóa bỏ vào trong hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1467)

Việc tạo ra cấu trúc 1467 được mô tả trong patent Mỹ số 61/541,780. Tóm lại, trình tự mã hoá HA từ cúm B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ được tạo dòng vào trong 2X35S/CPMV-HT/NOS bao gồm hệ khuếch đại BeYDV(m)+replicaza trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp thắt dựa trên PCR (Darveau et al. 1995, Methods in Neuroscience 26: 77-85). Trong vòng thứ nhất của PCR, mảnh chứa trình tự mã hoá HA B Wisconsin từ nt 1 đến nt 1062 được khuếch đại bằng cách sử dụng các đoạn mồi IF-HAB110.S1+3c (Fig.6A, SEQ ID NO: 49) và HAB110(PrL-).r (Fig.7A, SEQ ID NO: 55), bằng cách sử dụng HA B Wisconsin tổng hợp (Số truy cập Genbank JN993010) (Fig.6C, SEQ ID NO: 51) làm mẫu. Mảnh thứ hai, chứa trình tự mã hoá HA B Wisconsin từ nt 1120 đến nt 1755, được khuếch đại bằng cách sử dụng đoạn mồi HAB110(PrL-).c (Fig.7B, SEQ ID NO: 56) và IF-HAB110.s1-4r (Fig.6B, SEQ ID NO: 50), bằng cách sử dụng gen HA B Wisconsin tổng hợp (Số truy cập Genbank JN993010) (Fig.6C, SEQ ID NO: 51) làm mẫu. Sản phẩm PCR từ cả hai lần khuếch đại sau đó được trộn lần và sử dụng làm mẫu cho vòng khuếch đại thứ hai bằng cách sử dụng IF-HAB110.S1+3c (Fig.6A, SEQ ID NO: 49) và IF-HAB110.s1-4r (Fig.6B, SEQ ID NO: 50) làm đoạn mồi. Mảnh thu được (mã hoá HA B/Wisconsin/1/2010 Δa.a. 340-358 với cầu liên kết GG giữa các mảnh) được tạo dòng vào cátxét biểu hiện 2X35S/CPMV-HT/NOS bao gồm hệ khuếch đại BeYDV(m) bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 193 (Fig.6D) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion.

Cấu trúc số 193 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong cátxét biểu hiện được tạo thành chủ yếu từ CPMV-HT vào trong hệ khuếch đại BeYDV(m). Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên Fig.6E (SEQ ID NO: 52). Cấu trúc thu được được ký hiệu số 1467 (Fig.7C, SEQ ID NO: 57). Trình tự axit amin của HA từ cúm B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ được thể hiện trên Fig.7D (SEQ ID NO: 58). Phép biểu diễn plasmit 1467 được thể hiện trên Fig.7E.

Ví dụ 5.7: A-2X35S/CPMV-HT/PDISP-HA B Brisbane với vòng phân giải protein được xóa bỏ (Cấu trúc số 1039)

Việc tạo ra cấu trúc 1192 được mô tả trong patent Mỹ số 61/541,780. Tóm lại, trình tự mã hoá HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ được tạo dòng vào trong 2X35S/CPMV-HT/PDISP/NOS trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp thắt dựa trên PCR sau đây (Darveau et al., 1995, Methods in Neuroscience 26: 77-85). Trong vòng thứ nhất của PCR, mảnh chứa trình tự mã hoá HA B Brisbane từ nt 46 đến nt 1065 được khuếch đại bằng cách sử dụng gen HA B Brisbane tổng hợp (tương ứng với nt 34-1791 từ số truy cập Genebank FJ766840) làm mẫu. Mảnh thứ hai, chứa trình tự mã hoá HA B Brisbane từ nt 1123 đến nt 1758, được khuếch đại bằng cách sử dụng gen HA B Brisbane tổng hợp (tương ứng với nt 34-1791 từ Số truy cập Genbank FJ766840) làm mẫu. Sản phẩm PCR từ cả hai lần khuếch đại sau đó được trộn lần và sử dụng làm mẫu cho vòng khuếch đại thứ hai. Mảnh thu được (mã hoá HA B/Brisbane/60/2008 Δa.a. 356-374 với cầu liên kết GG giữa các mảnh) được tạo dòng trong khung với peptit tín hiệu alfalfa PDI trong cátxét biểu hiện 2X35S/CPMV-HT/NOS. Cấu trúc 1192 được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion.

Cấu trúc số 1192 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong khung với peptit tín hiệu alfalfa PDI trong cátxét biểu hiện dựa trên CPMV-HT. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA. Cấu trúc thu được được ký hiệu số 1039 (Fig.8B). Trình tự axit amin của PDISP-HA B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ được thể hiện trên Fig.5D (SEQ ID NO: 41). Phép biểu diễn plasmit 1039 được thể hiện trên Fig.8A (SEQ ID NO: 15).

Ví dụ 5.8: A-2X35S/CPMV-HT/H5 từ A/Indonesia/5/2005 với vị trí tách đột biến TETR (Cấu trúc số 676)

Trình tự mã hoá H5 từ A/Indonesia/5/2005 với vị trí tách đột biến TETR được tạo dòng vào trong 2X35S/CPMV-HT/NOS trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp thắt dựa trên PCR sau đây (Darveau et al., 1995, Methods in Neuroscience 26: 77-85). Trong vòng thứ nhất của PCR, mảnh chứa H5 từ trình tự mã hoá A/Indonesia/5/2005 từ nt 1 đến nt 1015 được khuếch đại bằng cách sử dụng đoạn mồi IF-H5A-I-05.s1+3c (Fig.1A, SEQ ID NO: 2) và MutCleavage-H5(Indo).r (Fig.23A, SEQ ID NO: 74), bằng cách sử dụng H5 tổng hợp từ A/Indonesia/5/2005 (Fig.1G, SEQ ID NO: 42) làm mẫu. Mảnh thứ hai, chứa H5 từ trình tự mã hoá A/Indonesia/5/2005 từ nt 1038 đến nt 1707, được khuếch đại bằng cách sử dụng đoạn mồi MutCleavage-H5(Indo).c (Fig.23B, SEQ ID NO : 75) và IF-H5dTm.r (Fig.1B, SEQ ID NO: 3), bằng cách sử dụng H5 tổng hợp từ A/Indonesia/5/2005 (Fig.1G, SEQ ID NO: 42) làm mẫu. Sản phẩm PCR từ cả hai lần khuếch đại sau đó được trộn lần và sử dụng làm mẫu cho vòng khuếch đại thứ hai bằng cách sử dụng IF-H5A-I-05.s1+3c (Fig.1A, SEQ ID NO: 2) và IF-H5dTm.r (Fig.1B, SEQ ID NO: 3) làm đoạn mồi. Mảnh thu được (mã hoá H5 từ A/Indonesia/5/2005 Δa.a. 339-346 với cầu liên kết TETR giữa các mảnh) được tạo dòng vào cátxét biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1191 (Fig.1D) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1191 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong khung trong cátxét biểu hiện dựa trên CPMV-HT. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên Fig.1D (SEQ ID NO: 4). Cấu trúc thu được được ký hiệu số 676 (Fig.23C, SEQ ID NO: 76). Trình tự axit amin của H5 từ thể đột biến vị trí tách A/Indonesia/5/2005 TETR được thể hiện trên Fig.23D (SEQ ID NO: 77). Phép biểu diễn theo sơ đồ của plasmit 676 được thể hiện trên Fig.23E.

Ví dụ 5.9: B-2X35S/CPMV-HT/H5 từ A/Indonesia/5/2005 với vị trí tách đột biến TETQ (Cấu trúc số 766)

Trình tự mã hoá H5 từ A/Indonesia/5/2005 với vị trí tách đột biến TETQ được tạo dòng vào trong 2X35S/CPMV-HT/ NOS trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp thắt dựa trên.PCR sau đây (Darveau et al., 1995, Methods in Neuroscience 26: 77-85). Trong vòng thứ nhất của PCR, mảnh chứa H5 từ trình tự mã hoá A/Indonesia/5/2005 từ nt 1 đến nt 1015 được khuếch đại bằng cách sử dụng đoạn mồi IF-H5A-I-05.s1+3c (Fig.1A, SEQ ID NO: 2) và H5I505_TETQ.r (Fig.24A, SEQ ID NO: 78), bằng cách sử dụng H5 tổng hợp từ A/Indonesia/5/2005 (Fig.1G, SEQ ID NO: 42) làm mẫu. Mảnh thứ hai, chứa H5 từ trình tự mã hoá A/Indonesia/5/2005 từ nt 1038 đến nt 1707, được khuếch đại bằng cách sử dụng đoạn mồi H5I505_TETQ.c (Fig.24B, SEQ ID NO : 79) và IF-H5dTm.r (Fig.1B, SEQ ID NO: 3), bằng cách sử dụng H5 tổng hợp từ A/Indonesia/5/2005 (Fig.1G, SEQ ID NO: 42) làm mẫu. Sản phẩm PCR từ cả hai lần khuếch đại sau đó được trộn lần và sử dụng làm mẫu cho vòng khuếch đại thứ hai bằng cách sử dụng IF-H5A-I-05.s1+3c (Fig.1A, SEQ ID NO: 2) và IF-H5dTm.r (Fig.1B, SEQ ID NO: 3) làm đoạn mồi. Mảnh thu được (mã hoá H5 from A/Indonesia/5/2005 Δa.a. 339-346 với cầu liên kết TETQ giữa các mảnh) được tạo dòng trong cátxét biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1191 (Fig.1D) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1191 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong khung trong cátxét biểu hiện dựa trên CPMV-HT. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên Fig.1D (SEQ ID NO: 4). Cấu trúc thu được được ký hiệu số 766 (Fig.24C, SEQ ID NO: 80). Trình tự axit amin của H5 từ thể đột biến vị trí tách A/Indonesia/5/2005 TETQ được thể hiện trên Fig.24D (SEQ ID NO: 81). Phép biểu diễn theo sơ đồ của plasmit 766 được thể hiện trên Fig.24E.

Ví dụ 5.10: C-2X35S/CPMV-HT/H5 từ A/Indonesia/5/2005 với vòng phân giải protein được xóa bỏ (Cấu trúc số 928)

Trình tự mã hoá H5 from A/Indonesia/5/2005 với vòng phân giải protein được xóa bỏ được tạo dòng vào trong 2X35S/CPMV-HT/ NOS trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp thất dựa trên PCR sau đây được thể hiện bởi Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Trong vòng thứ nhất của PCR, mảnh chứa H5 từ trình tự mã hoá A/Indonesia/5/2005 từ nt 1 đến nt 1011 được khuếch đại bằng cách sử dụng đoạn mồi IF-H5A-I-05.s1+3c (Fig.1A, SEQ ID NO: 2) và H5I505(PrL-).r (Fig.25A, SEQ ID NO: 82), bằng cách sử dụng H5 tổng hợp từ A/Indonesia/5/2005 (Fig.1G, SEQ ID NO: 42) làm mẫu. Mảnh thứ hai, chứa H5 từ trình tự mã hoá A/Indonesia/5/2005 từ nt 1075 đến nt 1707, được khuếch đại bằng cách sử dụng đoạn mồi H5I505(PrL-).c (Fig.25B, SEQ ID NO: 83) và IF-H5dTm.r (Fig.1B, SEQ ID NO: 3), bằng cách sử dụng H5 tổng hợp từ A/Indonesia/5/2005 (Fig.1G, SEQ ID NO: 42) làm mẫu. Sản phẩm PCR từ cả hai lần khuếch đại sau đó được trộn lần và sử dụng làm mẫu cho vòng khuếch đại thứ hai bằng cách sử dụng IF-H5A-I-05.s1+3c (Fig.1A, SEQ ID NO: 2) và IF-H5dTm.r (Fig.1B, SEQ ID NO: 3) làm đoạn mồi. Mảnh thu được (mã hoá H5 từ A/Indonesia/5/2005 Δa.a. 338-358 với cầu liên kết GG giữa các mảnh) được tạo dòng trong cátxét biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1191 (Fig.1D) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1191 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong khung trong cátxét biểu hiện dựa trên CPMV-HT. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên Fig.1D (SEQ ID NO: 4). Cấu trúc thu được được ký hiệu số 928 (Fig.25C, SEQ ID NO: 84). Trình tự axit amin của H5 từ A/Indonesia/5/2005 với vòng phân giải protein được xóa bỏ được thể hiện trên Fig.25D (SEQ ID NO: 85). Phép biểu diễn plasmit 928 được thể hiện trên Fig.25E.

Ví dụ 5.11 - F-2X35S/CPMV-HT/PDISP/HA B Brisbane/NOS (Cấu trúc số 1029)

Trình tự mã hoá HA từ cúm B/Brisbane/60/2008 được tạo dòng vào trong hệ biểu hiện 2X35S/CPMV-HT/PDISP/NOS trong plasmit chứa cátxét biểu hiện Plasto_pro/P19/Plasto_ter bằng cách sử dụng phương pháp dựa trên PCR sau đây. Mảnh chứa trình tự mã hoá HA B Brisbane mà không cần peptit tín hiệu kiểu hoang dại his được khuếch đại bằng cách sử dụng đoạn mồi IF-S2+S4-B Bris.c (Fig.30A, SEQ ID NO: 86) và IF-S1a4-B Bris.r (Fig.30B, SEQ ID NO: 87), bằng cách sử dụng gen HA B Brisbane tổng hợp (tương ứng với nt 34-1791 từ Số truy cập Genbank FJ766840) (Fig.30C, SEQ ID NO: 88) làm mẫu. Sản phẩm PCR được tạo dòng trong khung với peptit tín hiệu alfalfa PDI trong hệ biểu hiện 2X35S/CPMV-HT/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc 1192 được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 1192 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong khung với peptit tín hiệu alfalfa PDI trong cátxét biểu hiện dựa trên CPMV-HT. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN. Cấu trúc thu được được ký hiệu số 1029 (Fig.30D, SEQ ID NO: 89). Trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008 được thể hiện trên Fig.30E (SEQ ID NO: 90). Phép biểu diễn plasmit 1029 được thể hiện trên Fig.30F.

Ví dụ 5.12 - 2X35S/CPMV HT (cấu trúc số 1039) và HT+ (cấu trúc số 1829) đối với PDISP/HA B Brisbane (PrL-)

Trình tự mã hoá tương ứng với HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ (PrL-) trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/HA B Brisbane (PrL-; Fig.31A, SEQ ID NO: 91) được tạo dòng vào trong HT gốc và HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên PCR được mô tả trong các Ví dụ 5.7 và 5.11, nhưng với các đoạn mồi PCR cải biến được xác định một cách cụ thể đối với PDISP/HA B Brisbane (PrL-). Trình tự axit amin của HA B Brisbane (PrL-) chín được dung hợp với PDISP được thể

hiện trên Fig.31B (SEQ ID NO: 92). Phép biểu diễn của plasmit 1039 và 1829 được thể hiện trên Fig.8B và Fig.31D.

Ví dụ 5.13 - 2X35S/CPMV HT (cấu trúc số 1039) và 2X35S/CPMV160+ (cấu trúc số 1937) đối với PDISP/HA B Brisbane (PrL-)

Trình tự mã hoá tương ứng với HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein được xóa bỏ (PrL-) (xem đơn yêu cầu cấp patent tạm thời Mỹ số 61/806,227 nộp ngày 28 tháng 3 năm 2013, đối với thông tin bổ sung tham khảo: vùng vòng phân giải protein được xóa bỏ trong trình tự HA) trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/HA B Brisbane (PrL-)) (Fig.32A, SEQ ID NO: 93) được tạo dòng vào trong CPMV-HT và CPMV160+ gốc bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.7 và trong Ví dụ 5.11, nhưng với các đoạn mồi PCR cải biến được xác định một cách cụ thể đối với PDISP/HA B Brisbane (PrL-). Trình tự axit amin của HA B Brisbane (PrL-) chín được dung hợp với PDISP được thể hiện trên Fig.31B (SEQ ID NO: 92). Phép biểu diễn của plasmit 1039 và 1937 được thể hiện trên Fig.8B và Fig.32C.

Ví dụ 5.14 - 2X35S/CPMV HT (cấu trúc số 1067) và 2X35S/CPMV160+ (cấu trúc số 1977) đối với PDISP/HA B Brisbane (PrL-)+H1 California TMCT

Trình tự mã hoá hemagglutinin khảm tương ứng với ectodomain của HA từ cúm B/Brisbane/60/08 với vòng phân giải protein được xóa bỏ (PrL-) (xem đươn yêu cầu cấp patent tạm thời Mỹ số 61/806,227 nộp ngày 28 tháng 3 năm 2013, đối với thông tin bổ sung tham khảo: vùng vòng phân giải protein được xóa bỏ trong trình tự HA) được dung hợp với miền xuyên màng và đuôi tế bào chất (TMCT) của H1 từ cúm A/California/7/2009 và với peptit tín hiệu của alfalfa protein disulfua isomeraza (PDISP/HA B Brisbane (PrL-)+H1 California TMCT) (Fig.33A, SEQ ID NO: 95) được tạo dòng vào trong CPMV-HT và CPMV160+ gốc bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụs 5.7 và 5.11, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/HA B Brisbane (PrL-)+H1 California TMCT. Trình tự axit amin của HA B Brisbane (PrL-)+H1 California TMCT chín được

dung hợp với PDISP được thể hiện trên Fig.33B (SEQ ID NO: 96). Phép biểu diễn của plasmit 1067 và 1977 được thể hiện trên Fig.33C và Fig.33D.

Ví dụ 5.15 - 2X35S/CPMV HT (cấu trúc số 2072) và 2X35S/CPMV160+ (cấu trúc số 2050) đối với PDISP/HA B Massachussetts (PrL-)

Trình tự mã hoá tương ứng với HA từ cúm B/Massachussetts/2/2012 với vòng phân giải protein được xóa bỏ (PrL-) (xem đơn yêu cầu cấp patent tạm thời Mỹ số 61/806,227 nộp ngày 28 tháng 3 năm 2013 đối với các thông tin bổ sung tham khảo: vùng vòng phân giải protein được xóa bỏ trong trình tự HA) trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/HA B Massachussetts (PrL-)) (Fig.34A, SEQ ID NO: 97) được tạo dòng vào trong CPMV-HT và CPMV160+ gốc bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.7 và 5.11, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/HA B Massachussetts (PrL-). Trình tự axit amin HA B Massachussetts (PrL-) chín được dung hợp với PDISP được thể hiện trên Fig.34B (SEQ ID NO: 98). Phép biểu diễn của plasmit 2072 và 2050 được thể hiện trên Fig.34C và Fig.34D.

Ví dụ 5.16 - 2X35S/CPMV HT (cấu trúc số 2074) và 2X35S/CPMV160+ (cấu trúc số 2060) đối với PDISP/HA B Massachussetts (PrL-)+H1 California TMCT

Trình tự mã hoá hemagglutinin khảm tương ứng với ectodomain của HA từ cúm B/Massachussetts/2/2012 với vòng phân giải protein được xóa bỏ (PrL-) (xem đơn yêu cầu cấp patent tạm thời Mỹ số 61/806,227 nộp ngày 28 tháng 3 năm 2013 đối với các thông tin bổ sung tham khảo: vùng vòng phân giải protein được xóa bỏ trong trình tự HA) được dung hợp với miền xuyên màng và đuôi tế bào chất (TMCT) của H1 từ cúm A/California/7/2009 và với peptit tín hiệu của alfalfa protein disulfua isomeraza (PDISP/HA B Massachussetts (PrL-)+H1 California TMCT) (Fig.35A, SEQ ID NO: 99) được tạo dòng vào trong CPMV-HT và CPMV160+ gốc bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.7 và 5.11, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/HA B Massachussetts (PrL-)+H1 California TMCT. Trình tự axit amin của HA B Massachussetts (PrL-)+H1 California

TMCT chín được dung hợp với PDISP được thể hiện trên Fig.35B (SEQ ID NO: 100). Phép biểu diễn của plasmit 2074 và 2060 được thể hiện trên Fig.35C và Fig.35D.

Ví dụ 5.17 - 2X35S/CPMV HT (cấu trúc số 1445), 2X35S/CPMVHT+ (cấu trúc số 1820) và CPMV160+ (cấu trúc số 1975) đối với HA B Wisconsin (PrL-)

Trình tự mã hoá tương ứng với HA từ cúm B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ (PrL-) (xem đơn yêu cầu cấp patent tạm thời Mỹ số 61/806,227 nộp ngày 28 tháng 3 năm 2013 đối với các thông tin bổ sung tham khảo: vùng vòng phân giải protein được xóa bỏ trong trình tự HA) với peptit tín hiệu nguyên thể his (HA B Wisconsin (PrL-)) (Fig.36AA, SEQ ID NO: 101) được tạo dòng vào trong CPMV-HT, CPMVHT+, và CPMV160 gốc bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.7 và 5.11, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với HA B Wisconsin (PrL-). Trình tự axit amin của HA B Wisconsin (PrL-) với peptit tín hiệu nguyên thể his được thể hiện trên Fig.36B (SEQ ID NO: 102). Phép biểu diễn của plasmit 1445, 1820 và 1975 lần lượt được thể hiện trên Fig.36C, Fig.36D và Fig.36E.

Ví dụ 5.18 - 2X35S/CPMV HT (cấu trúc số 1454) và 2X35S/CPMV160+ (cấu trúc số 1893) đối với HA B Wisconsin (PrL-)+H1 California TMCT

Trình tự mã hoá hemagglutinin khảm tương ứng với ectodomain của HA từ cúm B/ Wisconsin /2/2012 với vòng phân giải protein được xóa bỏ (PrL-) (xem đơn yêu cầu cấp patent tạm thời Mỹ số 61/806,227 nộp ngày 28 tháng 3 năm 2013 đối với các thông tin bổ sung tham khảo: vùng vòng phân giải protein được xóa bỏ trong trình tự HA) được dung hợp với miền xuyên màng và đuôi tế bào chất (TMCT) của H1 từ cúm A/California/7/2009 với peptit tín hiệu nguyên thể của HA B Wisconsin (HA B Wisconsin (PrL-)+H1 California TMCT) (Fig.37A, SEQ ID NO: 103) được tạo dòng vào trong CPMV-HT và CPMV160+ gốc bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.7 và 5.11, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với HA B Wisconsin (PrL-)+H1 California TMCT. Trình tự axit amin của HA B Wisconsin (PrL-)+H1 California TMCT được thể hiện trên Fig.37 (SEQ ID NO: 104). Phép biểu diễn của plasmit 1454 và 1893 được thể hiện trên Fig.37C và Fig.37D.

Ví dụ 5.19: 2X35S/CPMV HT (cấu trúc số 1067) và HT+ (cấu trúc số 1875) đối với PDISP/HA B Brisbane (PrL-)+H1 California TMCT

Trình tự mã hoá hemagglutinin khảm tương ứng với ectodomain của HA từ cúm B/Brisbane/60/08 với vòng phân giải protein được xóa bỏ (PrL-) được dung hợp với miền xuyên màng và đuôi tế bào chất (TMCT) của H1 từ cúm A/California/7/2009 và với peptit tín hiệu của alfalfa protein disulfua isomeraza (PDISP/HA B Brisbane (PrL-)+H1 California TMCT) (Fig.38A, SEQ ID NO: 105) được tạo dòng vào trong HT gốc và HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.26, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/HA B Brisbane (PrL-)+H1 California TMCT. Trình tự axit amin của HA B Brisbane (PrL-)+H1 California TMCT chín được dung hợp với PDISP được thể hiện trên Fig.38B (SEQ ID NO: 106). Phép biểu diễn của plasmit 1067 và 1875 được thể hiện trên Fig.33C và 39C.

Ví dụ 5.20: 2X35S/CPMV HT (cấu trúc số 2072) và HT+ (cấu trúc số 2052) đối với PDISP/HA B Massachussetts (PrL-)

Trình tự mã hoá tương ứng với HA từ cúm B/Massachusetts/2/2012 với vòng phân giải protein được xóa bỏ (PrL-) trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/HA B Massachussetts (PrL-)) (Fig.39A, SEQ ID NO: 107) được tạo dòng vào trong HT gốc và HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.26, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/HA B Massachussetts (PrL-). Trình tự axit amin của HA B Massachussetts (PrL-) chín được dung hợp với PDISP được thể hiện trên Fig.39B (SEQ ID NO: 108). Phép biểu diễn của plasmit 2072 và 2052 được thể hiện trên Fig.34C và Fig.39C.

Ví dụ 5.21: 2X35S/CPMV HT (cấu trúc số 2074) và HT+ (cấu trúc số 2062) đối với PDISP/HA B Massachussetts (PrL-)+H1 California TMCT

Trình tự mã hoá hemagglutinin khảm tương ứng với ectodomain của HA từ cúm B/Massachussetts/2/2012 với vòng phân giải protein được xóa bỏ (PrL-) được dung hợp với miền xuyên màng và đuôi tế bào chất (TMCT) của H1 từ cúm A/California/7/2009 và với peptit tín hiệu của alfalfa protein disulfua isomeraza (PDISP/HA B Massachussetts

(PrL-)+H1 California TMCT) (Fig.40A, SEQ ID NO: 109) được tạo dòng vào trong HT gốc và HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.26, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/HA B Massachussetts (PrL-)+H1 California TMCT. Trình tự axit amin của HA B Massachussetts (PrL-)+H1 California TMCT chín được dung hợp với PDISP được thể hiện trên Fig.40B (SEQ ID NO: 110). Phép biểu diễn của plasmit 2074 và 2062 được thể hiện trên Fig.35C và Fig.40C.

Ví dụ 5.22: 2X35S/CPMV HT (cấu trúc số 1445) và HT+ (cấu trúc số 1839) đối với HA B Wisconsin (PrL-)

Trình tự mã hoá tương ứng với HA từ cúm B/Wisconsin/1/2010 với vòng phân giải protein được xóa bỏ (PrL-) với peptit tín hiệu nguyên thể his (HA B Wisconsin (PrL-)) (Fig.41A, SEQ ID NO: 111) được tạo dòng vào trong HT gốc và HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.26, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với HA B Wisconsin (PrL-). Trình tự axit amin của HA B Wisconsin (PrL-) với peptit tín hiệu nguyên thể his được thể hiện trên Fig.41B (SEQ ID NO: 112). Phép biểu diễn của plasmit 1445 và 1839 được thể hiện trên Fig.36C và Fig.41C.

Ví dụ 5.23: 2X35S/CPMV HT (cấu trúc số 1454) và HT+ (cấu trúc số 1860) đối với HA B Wisconsin (PrL-)+H1 California TMCT

Trình tự mã hoá hemagglutinin khảm tương ứng với ectodomain của HA từ cúm B/ Wisconsin /2/2012 với vòng phân giải protein được xóa bỏ (PrL-) được dung hợp với miền xuyên màng và đuôi tế bào chất (TMCT) của H1 từ cúm A/California/7/2009 với peptit tín hiệu nguyên thể của HA B Wisconsin (HA B Wisconsin (PrL-)+H1 California TMCT) (Fig.42A, SEQ ID NO: 113) được tạo dòng vào trong HT gốc và HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.26 nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với HA B Wisconsin (PrL-)+H1 California TMCT. Trình tự axit amin của HA B Wisconsin (PrL-)+H1 California TMCT được thể hiện trên Fig.42B (SEQ ID NO: 114). Phép biểu diễn của plasmit 1454 và 1860 được thể hiện trên Fig.37C và Fig.42C.

Ví dụ 5.24 - 2X35S/CPMV HT (cấu trúc số 489), 2X35S/CPMV160+ (cấu trúc số 1880) và 2X35S/CPMV160 (cấu trúc số 1885) đối với H5 Indonesia

Trình tự mã hoá tương ứng với H5 nguyên thể từ cúm A/Indonesia/5/2005 (Fig.43A, SEQ ID NO: 115) được tạo dòng vào trong CPMV-HT, CPMV160+ và CPMV160 gốc bằng cách sử dụng cùng phương pháp dựa trên PCR như được mô tả trong Ví dụ 5.25 nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với H5 Indonesia. Trình tự axit amin của H5 nguyên thể từ cúm A/Indonesia/5/2005 được thể hiện trên Fig.43B (SEQ ID NO: 116). Phép biểu diễn của plasmit 489 được thể hiện trên Fig.43C.

Ví dụ 5.25 - 2X35S/CPMV160+/PDISP/H3 Victoria/ NOS (Cấu trúc số 1800)

Trình tự mã hoá H3 từ cúm A/Victoria/361/2011 trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H3 Victoria) được tạo dòng vào trong hệ biểu hiện 2X35S/CPMV160+/NOS (CPMV160+) bằng cách sử dụng phương pháp dựa trên PCR sau đây. Mảnh chứa trình tự mã hoá PDISP/H3 Victoria được khuếch đại bằng cách sử dụng đoạn mồi IF**(SacII)-PDI.s1+4c (Fig.44A, SEQ ID NO: 117) và IF-H3V36111.s1-4r (Fig.44B, SEQ ID NO: 118), bằng cách sử dụng trình tự PDISP/H3 Victoria (Fig.44C, SEQ ID NO: 119) làm mẫu. Sản phẩm PCR được tạo dòng trong hệ biểu hiện 2X35S/CPMV160+/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc số 2171 (Fig.44D) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 2171 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong cátxét biểu hiện dựa trên CPMV160+. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên Fig.44E (SEQ ID NO: 120). Cấu trúc thu được được ký hiệu số 1800 (Fig.44F, SEQ ID NO: 121). Trình tự axit amin của H3 chín từ cúm A/Victoria/361/2011 được dung hợp với PDISP được thể hiện trên Fig.44G (SEQ ID NO: 122). Phép biểu diễn plasmit 1800 được thể hiện trên Fig.44H.

Ví dụ 5.26: 2X35S/CPMV-HT+/ PDISP/H3 Victoria/ NOS (Cấu trúc số 1819)

Trình tự mã hoá H3 từ cúm A/Victoria/361/2011 trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H3 Victoria) được tạo dòng vào trong hệ biểu hiện 2X35S-CPMV-HT+/NOS bằng cách sử dụng phương pháp dựa trên PCR sau đây. Mảnh chứa PDISP/H3 Victoria trình tự mã hoá được khuếch đại bằng cách sử dụng đoạn mồi IF(SacII)-Kozac_PDI.c (Fig.45A, SEQ ID NO: 123) và IF-H3V36111.s1-4r (Fig.45B, SEQ ID NO: 124), bằng cách sự dụng trình tự PDISP/H3 Victoria (Fig.44C, SEQ ID NO: 119) làm mẫu. Sản phẩm PCR được tạo dòng trong hệ biểu hiện 2X35S/CPMV-HT+/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc số 2181 (Fig.45D) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 2181 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong cátxét biểu hiện dựa trên CPMV-HT+. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên Fig.45E (SEQ ID NO: 125). Cấu trúc thu được được ký hiệu số 1819 (Fig.45E, SEQ ID NO: 126). Trình tự axit amin của H3 chín từ cúm A/Victoria/361/2011 được dung hợp với PDISP được thể hiện trên Fig.44G (SEQ ID NO: 122). Phép biểu diễn plasmit 1819 được thể hiện trên Fig.45F.

Ví dụ 5.27 2X35S/CPMV HT+/ PDISP/H2 Singapore/ NOS (Cấu trúc số 2220)

Trình tự mã hoá H2 từ cúm A/Singapore/1/1957 trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H2 Singapore) được tạo dòng vào trong hệ biểu hiện 2X35S/CPMV HT+/NOS bằng cách sử dụng phương pháp dựa trên PCR sau đây. Mảnh chứa trình tự mã hoá PDISP/H2 Singapore được khuếch đại bằng cách sử dụng đoạn mồi IF(SacII)-Kozac_PDI.c (được mô tả đối với cấu trúc 1819 trong Ví dụ 5.26) và IF**-H2S157.s1-6r (Fig.48A, SEQ ID NO: 127), bằng cách sử dụng trình tự PDISP/H2 Singapore (Fig.48B, SEQ ID NO : 128) làm mẫu. Sản phẩm PCR được tạo dòng trong hệ biểu hiện 2X35S/CPMV HT+/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc số 2181 (được mô tả đối

với cấu trúc 1819 trong Ví dụ 5.26) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 2181 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong cátxét biểu hiện dựa trên CPMV HT+. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên Fig.45D. Cấu trúc thu được được ký hiệu số 2220 (Fig.48C, SEQ ID NO: 129). Trình tự axit amin của H2 chín từ cúm A/Singapore/1/1957 được dung hợp với PDISP được thể hiện trên Fig.48D (SEQ ID NO: 130). Phép biểu diễn plasmit 2220 được thể hiện trên Fig.48E.

Ví dụ 5.28 2X35S/CPMV HT+/ PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ/ NOS (Cấu trúc số 2221)

Trình tự mã hoá H2 từ cúm A/Singapore/1/1957 với vòng phân giải protein được xóa bỏ trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ) được tạo dòng vào trong hệ biểu hiện 2X35S/CPMV HT+/NOS bằng cách sử dụng phương pháp thất dựa trên PCR sau đây được đại diện bởi Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Trong vòng thứ nhất của PCR, mảnh chứa trình tự mã hoá H2 từ cúm A/Singapore/1/1957 từ nt 1 đến nt 1032 được khuếch đại đối với các đoạn mồi IF(SacII)-Kozac_PDI.c (được mô tả đối với cấu trúc 1819 trong Ví dụ 5.26) và H2S157(Prl-).r (Fig.49A, SEQ ID NO: 131), bằng cách sử dụng trình tự PDISP/H2 Singapore (Fig.48B, SEQ ID NO: 128) làm mẫu. Mảnh thứ hai, chứa trình tự mã hoá H2 từ cúm A/Singapore/1/1957 từ nt 1084 đến nt 1716, được khuếch đại với các đoạn mồi H2S157(Prl-).c (Fig.49B, SEQ ID NO : 132) và IF**-H2S157.s1-6r (Fig.48A, SEQ ID NO: 127) bằng cách sử dụng trình tự PDISP/H2 Singapore (Fig.48B, SEQ ID NO :128) làm mẫu. Sản phẩm PCR từ cả hai lần khuếch đại sau đó được trộn lần và sử dụng làm mẫu cho vòng khuếch đại thứ hai bằng cách sử dụng IF(SacII)-Kozac_PDI.c (được mô tả đối với cấu trúc 1819 trong Ví dụ 5.26) và IF**-H2S157.s1-6r (Fig.48A, SEQ ID NO: 127) làm đoạn mồi. Sản phẩm PCR (bao gồm trình tự mã hoá PDISP/H2 Singapore với aa

321 đến 337 được thay thế bằng cầu liên kết GG) được tạo dòng vào hệ biểu hiện 2X35S/CPMV HT+/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc số 2181 (được mô tả đối với cấu trúc 1819 trong Ví dụ 5.26) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 2181 là plasmit nhận được dự định để tạo dòng gen "In Fusion" quan tâm trong cátxét biểu hiện dựa trên CPMV HT+. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên Fig.45D (được mô tả đối với cấu trúc 1819, trong Ví dụ 5.26). Cấu trúc thu được được ký hiệu số 2221 (Fig.49C, SEQ ID NO: 133). Trình tự axit amin của H2 chín từ cúm A/Singapore/1/1957 với vòng phân giải protein được xóa bỏ được dung hợp với PDISP được thể hiện trên Fig.49D (SEQ ID NO: 134). Phép biểu diễn plasmit 2221 được thể hiện trên Fig.49E.

Ví dụ 5.29 PDISP/H2 Singapore (Cấu trúc số 2222) và PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ (Cấu trúc số 2223) trong hệ biểu hiện 2X35S/CPMV 160+/NOS

Các trình tự mã hoá H2 từ cúm A/Singapore/1/1957 có hoặc không có vòng phân giải protein trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H2 Singapore và PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ) được tạo dòng vào trong hệ biểu hiện 2X35S/CPMV 160+/NOS bằng cách sử dụng cùng phương pháp dựa trên PCR như cấu trúc 2220 và 2221, một cách tương ứng, nhưng bằng cách sử dụng đoạn mồi thuận cải biến IF**(SacII)-PDI.s1+4c (được mô tả đối với cấu trúc 1800 trong Ví dụ 5.25) để khuếch đại và plasmit nhận khác nhau. Các sản phẩm PCR thu được được tạo dòng trong hệ biểu hiện 2X35S/CPMV 160+/NOS bằng cách sử dụng hệ tạo dòng In-Fusion (Clontech, Mountain View, CA). Cấu trúc số 2171 (được mô tả đối với cấu trúc 1800 trong Ví dụ 5.25) được tiêu hoá bằng enzym giới hạn SacII và StuI và plasmit tuyến tính hoá được sử dụng cho phản ứng lắp ráp In-Fusion. Cấu trúc số 2171 là plasmit nhận được dự định để tạo dòng gen "In Fusion"

quan tâm trong cátxét biểu hiện dựa trên CPMV 160+. Nó cũng hợp nhất cấu trúc gen để đồng biểu hiện gen ức chế gây câm TBSV P19 trong đoạn khởi đầu và đoạn kết thúc gen alfalfa Plastocyanin. Khung chính là plasmit nhị phân pCAMBIA và trình tự từ phía bên trái sang bên phải đường biên t-ADN được thể hiện trên (được mô tả đối với cấu trúc 1800 trong Ví dụ 5.25). Các cấu trúc thu được được ký hiệu số 2222 đối với PDISP/H2 Singapore (Fig.50A, SEQ ID NO: 135) và 2223 đối với PDISP/H2 Singapore với vòng phân giải protein được xóa bỏ (Fig.50B, SEQ ID NO: 136). Phép biểu diễn của plasmit 2222 và 2223 được thể hiện trên Fig.50C và 50D một cách tương ứng.

Ví dụ 5.30 2X35S/CPMV HT+ (cấu trúc số 2019) và 160+ (cấu trúc số 2139) đối với PDISP/H3 Perth

Trình tự mã hoá tương ứng với H3 từ cúm A/Perth/16/2009 trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H3 Perth) (Fig.51A, SEQ ID NO: 137) được tạo dòng vào trong CPMV HT+ và 160+ cải biến bằng cách sử dụng cùng phương pháp đựa trên In Fusion như cấu trúc 2220 và 2222, một cách tương ứng, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/H3 Perth (Fig.51B, SEQ ID NO: 138). Trình tự axit amin của H3 chín từ cúm A/Perth/16/2009 được dung hợp với PDISP được thể hiện trên Fig.51C (SEQ ID NO: 139). Phép biểu diễn của plasmit 2019 và 2139 được thể hiện trên Fig.51D và Fig.51E.

Ví dụ 5.31 2X35S/CPMV HT+ (cấu trúc số 2039) và 160+ (cấu trúc số 2159) đối với PDISP/H3 Perth với vòng phân giải protein được xóa bỏ

Trình tự mã hoá tương ứng với H3 từ cúm A/Perth/16/2009 với vòng phân giải protein được xóa bỏ trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H3 Perth với vòng phân giải protein được xóa bỏ) (Fig.52, SEQ ID NO: 140) được tạo dòng vào trong CPMV HT+ và 160+ cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như cấu trúc 2221 và 2223, một cách tương ứng, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/H3 Perth với vòng phân giải protein được xóa bỏ (Fig.51B (SEQ ID NO: 138), 52B (SEQ ID NO: 141) và 53C (SEQ ID NO: 142). Trình tự axit amin H3 chín từ cúm

A/Perth/16/2009 với vòng phân giải protein được xóa bỏ được dung hợp với PDISP được thể hiện trên Fig.52D (SEQ ID NO: 143). Phép biểu diễn của plasmit 2039 và 2159 được thể hiện trên Fig.52E và Fig.52F.

Ví dụ 5.32 2X35S/CPMV HT+ (cấu trúc số 2230) và 160+ (cấu trúc số 2250) đối với PDISP/H3 Victoria với vòng phân giải protein được xóa bỏ

Trình tự mã hoá tương ứng với H3 từ cúm A/Victoria/361/2011 với vòng phân giải protein được xóa bỏ trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H3 Victoria với vòng phân giải protein được xóa bỏ) (Fig.53, SEQ ID NO: 144) được tạo dòng vào trong CPMV HT+ và 160+ cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như cấu trúc 2221 và 2223 (xem ví dụs 5.28 và 5.29), một cách tương ứng, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/H3 Victoria với vòng phân giải protein được xóa bỏ (Fig.53B (SEQ ID NO: 145) và 53C (SEQ ID NO: 146). Trình tự axit amin của H3 chín từ cúm A/Victoria/361/2011 với vòng phân giải protein được xóa bỏ được dung hợp với PDISP được thể hiện trên Fig.53D (SEQ ID NO: 147). Phép biểu diễn của plasmit 2230 và 2250 được thể hiện trên Fig.53E và Fig.53F.

Ví dụ 5.33 2X35S/CPMV HT+/PDISP/H7 Hangzhou/NOS (cấu trúc số 2142)

Trình tự mã hoá tương ứng với H7 từ cúm A/Hangzhou/1/2013 trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H7 Hangzhou) (Fig.54A, SEQ ID NO: 148) được tạo dòng vào trong CPMV HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như cấu trúc 2220 nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/H7 Hangzhou (Fig.54B, SEQ ID NO: 149). Trình tự axit amin của H7 chín từ cúm A/Hangzhou/1/2013 được dung hợp với PDISP được thể hiện trên Fig.54C (SEQ ID NO: 150). Phép biểu diễn của plasmit 2142 được thể hiện trên Fig.54E.

Ví dụ 5.34 2X35S/CPMV HT+/PDISP/H7 Hangzhou với vòng phân giải protein được xóa bỏ/NOS (cấu trúc số 2152)

Trình tự mã hoá tương ứng với H7 từ cúm A/Hangzhou/1/2013 với vòng phân giải protein được xóa bỏ trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/ H7 Hangzhou với vòng phân giải protein được xóa bỏ) (Fig.55A, SEQ ID NO: 151) được tạo dòng vào trong CPMV HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như cấu trúc 2221 nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/ H7 Hangzhou với vòng phân giải protein được xóa bỏ (Fig.54B (SEQ ID NO: 149), Fig.55B (SEQ ID NO: 152) và Fig.55C (SEQ ID NO: 153)). Trình tự axit amin của H7 chín từ cúm A/Hangzhou/1/2013 với vòng phân giải protein được xóa bỏ được dung hợp với PDISP được thể hiện trên Fig.55D (SEQ ID NO: 154). Phép biểu diễn của plasmit 2152 được thể hiện trên Fig.55E.

Ví dụ 5.35 2X35S/CPMV HT+ (cấu trúc số 2224) và 160+ (cấu trúc số 2226) đối với PDISP/H9 Hong Kong

Trình tự mã hoá tương ứng với H9 từ cúm A/Hong Kong/1073/1999 trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H9 Hong Kong) (Fig.56A, SEQ ID NO: 155) được tạo dòng vào trong CPMV HT+ và 160+ cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như cấu trúc 2220 và 2222, một cách tương ứng, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/H9 Hong Kong (Fig.56B, SEQ ID NO: 156). Trình tự axit amin của H9 chín từ cúm A/Hong Kong/1073/1999 được dung hợp với PDISP được thể hiện trên Fig.56C (SEQ ID NO: 157). Phép biểu diễn của plasmit 2224 và 2226 được thể hiện trên Fig.56D và Fig.56E.

Ví dụ 5.36 2X35S/CPMV HT+ (cấu trúc số 2225) và 160+ (cấu trúc số 2227) đối với PDISP/H9 Hong Kong với vòng phân giải protein được xóa bỏ

Trình tự mã hoá tương ứng với H9 từ cúm A/Hong Kong/1073/1999 với vòng phân giải protein được xóa bỏ trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/H9 Hong Kong với vòng phân giải protein được xóa bỏ) (Fig.57A, SEQ ID NO: 158) được tạo dòng vào trong CPMV HT+ và 160+

cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như cấu trúc 2221 và 2223, một cách tương ứng, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/H9 Hong Kong với vòng phân giải protein được xóa bỏ (Fig.56B (SEQ ID NO: 156), Fig.57B (SEQ ID NO: 159) và Fig.57C (SEQ ID NO: 160)). Trình tự axit amin của H9 chín từ cúm A/Hong Kong/1073/1999 với vòng phân giải protein được xóa bỏ được dung hợp với PDISP được thể hiện trên Fig.57D (SEQ ID NO: 161). Phép biểu diễn của plasmit 2225 và 2227 được thể hiện trên Fig.57E và Fig.57F.

Ví dụ 5.37 2X35S/CPMV 160+/PDISP/HA B Malaysia/NOS (cấu trúc số 2013)

Trình tự mã hoá tương ứng với HA từ cúm B/Malaysia/2506/2004 trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/HA B Malaysia) (Fig.58A, SEQ ID NO: 162) được tạo dòng vào trong CPMV 160+ cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như cấu trúc 2222 nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/HA B Malaysia (Fig.58B, SEQ ID NO: 163). Trình tự axit amin của HA chín từ cúm B/Malaysia/2506/2004 được dung hợp với PDISP được thể hiện trên Fig.58C (SEQ ID NO: 164). Phép biểu diễn của plasmit 2013 được thể hiện trên Fig.58D.

Ví dụ 5.38 2X35S/CPMV 160+/PDISP/HA B Malaysia với vòng phân giải protein được xóa bỏ/NOS (cấu trúc số 2014)

Trình tự mã hoá tương ứng với HA từ cúm B/Malaysia/2506/2004 với vòng phân giải protein được xóa bỏ trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/HA B Malaysia với vòng phân giải protein được xóa bỏ) (Fig.59A, SEQ ID NO: 165) được tạo dòng vào trong CPMV 160+ cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như cấu trúc 2223 nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/HA B Malaysia với vòng phân giải protein được xóa bỏ (Fig.58B (SEQ ID NO: 163), Fig.59B (SEQ ID NO: 166), Fig.59C (SEQ ID NO: 167). Trình tự axit amin của HA chín từ cúm B/Malaysia/2506/2004 với vòng phân giải protein được xóa bỏ được dung hợp với PDISP

được thể hiện trên Fig.59D (SEQ ID NO: 168). Phép biểu diễn của plasmit 2014 được thể hiện trên Fig.59E.

Ví dụ 5.39 2X35S/CPMV HT (cấu trúc số 2070), HT+ (cấu trúc số 2080) và 160+ (-Mprot) (cấu trúc số 2090) đối với PDISP/HA B Massachusetts

Trình tự mã hoá tương ứng với HA từ cúm B/Massachusetts/2/2012 trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/ HA B Massachusetts) (Fig.60A, SEQ ID NO: 169) được tạo dòng vào trong HT gốc, HT+ cải biến và 160+ bằng cách sử dụng cùng phương pháp dựa trên In Fusion như cấu trúc 2072, 2220 và 2222, một cách tương ứng, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/ HA B Massachusetts. Trình tự axit amin của HA chín từ cúm B/Massachusetts/2/2012 được dung hợp với PDISP được thể hiện trên Fig.60B (SEQ ID NO: 170). Phép biểu diễn của plasmit 2070, 2080 và 2090 được thể hiện trên Fig.60C, Fig.60D và Fig.60E.

Ví dụ 5.40 2X35S/CPMV HT+ (cấu trúc số 2102), HT+ với BeYDV (cấu trúc số 2104) đối với PDISP/HA B Florida với vòng phân giải protein được xóa bỏ

Trình tự mã hoá tương ứng với HA từ cúm B/Florida với vòng phân giải protein được xóa bỏ và trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/ HA B Florida) (Fig.61D, SEQ ID NO: 193) được tạo dòng vào trong HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như được mô tả trên đây, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/ HA B/Florida (xem Fig.61A (SEQ ID No:190), Fig.61B (SEQ ID NO:191) và Fig.61C (SEQ ID NO: 192). Trình tự nucleotit của cátxét biểu hiện thu được 2102 được nêu trên Fig.61F (SEQ ID NO: 195). Tương tự, trình tự mã hoá tương ứng với HA từ cúm B/Florida với vòng phân giải protein được xóa bỏ và trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza được tạo dòng vào trong HT+ cải biến cùng với yếu tố khuếch đại BeYDV. Trình tự nucleotit của cátxét biểu hiện thu được 2104 được nêu trên Fig.61F (SEQ ID NO: 196). Trình tự axit amin của HA chín từ cúm B/Florida với vòng phân giải protein được xóa bỏ được dung

hợp với PDISP được thể hiện trên Fig.61E (SEQ ID NO: 194). Phép biểu diễn của plasmit 2102 và 2104 được thể hiện trên Fig.61G và Fig.61I.

Ví dụ 5.41 2X35S/CPMV HT+ (cấu trúc số 2106), HT+ với BeYDV (cấu trúc số 2108) đối với PDISP/HA B Florida +H1 California TMCT với vòng phân giải protein được xóa bỏ

Trình tự mã hoá tương ứng với HA từ cúm B/Florida +H1 California TMCT với vòng phân giải protein được xóa bỏ và trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza (PDISP/ HA B Florida +H1 California TMCT) (Fig.62B, SEQ ID NO: 198) được tạo dòng vào trong HT+ cải biến bằng cách sử dụng cùng phương pháp dựa trên In Fusion như được mô tả trên đây, nhưng với đoạn mồi PCR cải biến được ký hiệu một cách cụ thể đối với PDISP/ HA B/Florida+H1 California TMCT (xem Fig.61A (SEQ ID No:197). Trình tự nucleotit của cátxét biểu hiện thu được 2106 được nêu trên Fig.62D (SEQ ID NO: 200). Tương tự, trình tự mã hoá tương ứng với HA từ cúm B/Florida +H1 California TMCT với vòng phân giải protein được xóa bỏ và trong đó peptit tín hiệu nguyên thể đã được thay thế bằng peptit của alfalfa protein disulfua isomeraza được tạo dòng vào trong HT+ cải biến cùng với yếu tố khuếch đại BeYDV. Trình tự nucleotit của cátxét biểu hiện thu được 2108 được nêu trên Fig.62F (SEQ ID NO: 201). Trình tự axit amin của HA chín từ cúm B/Florida+H1 California TMCT với vòng phân giải protein được xóa bỏ được dung hợp với PDISP được thể hiện trên Fig.62C (SEQ ID NO: 199). Phép biểu diễn của plasmit 2106 và 2108 được thể hiện trên Fig.62E và Fig.62G.

Tất cả các tài liệu trích dẫn được đưa ra trong bản mô tả này để tham khảo.

Sáng chế đã được mô tả xét đến một hoặc nhiều phương án. Tuy nhiên, sẽ rõ ràng đối với chuyên gia trong lĩnh vực này rằng nhiều biến đổi và cải biến có thể được tạo ra mà không tách rời phạm vi của sáng chế như được xác định trong các điểm yêu cầu bảo hộ.

YÊU CẦU BẢO HỘ

1. Axit nucleic chứa vùng điều hoà có hoạt tính ở thực vật và vùng tăng cường sự biểu hiện có hoạt tính ở thực vật, vùng điều hòa và vùng tăng cường sự biểu hiện được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá hemagglutinin (HA) cúm được cải biến được chọn từ HA cúm typ B và HA cúm typ A typ phụ H3 và H7, trong đó HA cúm typ B không bao gồm chủng B/Brisbane/60/2008 hoặc chủng B/Wisconsin/1/2010, HA được cải biến này bao gồm vòng phân giải protein được xóa bỏ hoàn toàn giữa các dưới đơn vị HA1 và HA2, vòng phân giải protein được xóa bỏ hoàn toàn này chứa vị trí phân cắt monobazơ hoặc đa bazơ, và trong đó axit nucleic này không bao gồm yếu tố khuếch đại virut vàng lùn đậu (Bean Yellow Dwarf Virus - BeYDV).

2. Axit nucleic theo điểm 1, trong đó vùng tăng cường sự biểu hiện được chọn từ nhóm gồm có CPMVX, CPMVX+, CPMV HT+, CPMV HT+[WT115] và CPMV HT+[511], trong đó:

CPMVX chứa X nucleotit của SEQ ID NO:93, trong đó X=160, 155, 150, hoặc 114 của SEQ ID NO:93, hoặc trình tự mà có độ đồng nhất về mặt trình tự từ 80% đến 100% với CPMVX, trong đó X=160, 155, 150, hoặc 114 của SEQ ID NO:93,

CPMVX+ chứa X nucleotit của SEQ ID NO:93, trong đó X=160, 155, 150, hoặc 114 của SEQ ID NO:93, hoặc trình tự mà có độ đồng nhất về mặt trình tự từ 80% đến 100% với CPMVX, trong đó X=160, 155, 150, hoặc 114 của SEQ ID NO:93, và trình tự nhồi chứa từ 1-100 nucleotit được dung hợp với đầu 3' của trình tự CMPVX,

CPMV HT+ chứa vùng chưa dịch mã (untranslated region - UTR) đầu 5' comovirus và trình tự nhồi được cải biến, kéo dài hoặc cắt cụt,

CPMV HT+[WT115] chứa trình tự SEQ ID NO:189, và

CPMV HT+[511] chứa trình tự SEQ ID NO:188.

3. Axit nucleic theo điểm 1 hoặc 2, trong đó vùng tăng cường sự biểu hiện không phải là CPMV HT.

4. Axit nucleic theo điểm 1, 2 hoặc 3, trong đó HA được cải biến được giữ ở dạng tiền chất HA0.

5. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 4, trong đó vòng phân giải protein này được thay thế hoàn toàn bằng trình tự liên kết.

6. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 5, trong đó trình tự nucleotit mã hoá hemagglutinin (HA) protein cúm có ít nhất 70% trình tự giống với trình tự axit amin được chọn từ nhóm gồm có các axit amin từ 25 đến 576 của các SEQ ID NO: 98 và 100, các axit amin từ 25 đến 559 của SEQ ID NO: 143, các axit amin từ 25 đến 559 của SEQ ID NO: 147, các axit amin từ 25 đến 549 của SEQ ID NO: 154, các axit amin từ 25 đến 577 của SEQ ID NO: 168, và các axit amin từ 25 đến 576 của các SEQ ID NO: 194 và 199.

7. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 6, trong đó HA được cải biến chứa peptit tín hiệu HA nguyên thể hoặc peptit tín hiệu HA không nguyên thể.

8. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 6, trong đó trình tự nucleotit mã hoá HA được cải biến bao gồm trình tự nucleotit khảm mã hoá, theo dãy, HA ectodomain được cải biến bao gồm vòng phân giải protein được xóa bỏ hoàn toàn, miền xuyên màng cúm, và đuôi tế bào chất, trong đó HA ectodomain được cải biến là từ chủng cúm thứ nhất và miền xuyên màng và đuôi tế bào chất là từ chủng cúm thứ hai.

9. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 8, trong đó HA được cải biến là HA cúm typ A typ phụ H3 và vòng phân giải protein được xóa bỏ hoàn toàn gồm trình tự SEQ ID NO:48.

10. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 8, trong đó HA được cải biến là HA cúm typ B và vòng phân giải protein được xóa bỏ hoàn toàn gồm trình tự SEQ ID NO:59.

11. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 8, trong đó HA được cải biến là HA cúm typ A typ phụ H7 và vòng phân giải protein được xóa bỏ hoàn toàn gồm trình tự gồm các axit amin từ 340 đến 357 của SEQ ID NO:150.

12. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 8, trong đó vị trí phân cắt monobazơ được nhận biết bằng tryptaza Clara hoặc trypsin/chymotrypsin.

13. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 8, trong đó vị trí phân cắt đa bazơ được nhận biết bằng furin hoặc subtilisin.

14. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 8, trong đó vị trí phân cắt chứa trình tự axit amin KER, PAK hoặc tổ hợp của chúng đối với HA cúm typ B.

15. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 8, trong đó vị trí phân cắt chứa trình tự axit amin Q/E-X-R, PEK hoặc tổ hợp của chúng đối với HA cúm typ A typ phụ H3.

16. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 8, trong đó vị trí phân cắt chứa trình tự axit amin R-X-R/K-R, RKKR hoặc tổ hợp của chúng đối với HA cúm typ A typ phụ H7.

17. Phương pháp tạo ra hạt tương tự virut (virus like particle - VLP) cúm ở thực vật, phương pháp này bao gồm các bước:

a) đưa axit nucleic vào trong thực vật hoặc phần của thực vật, hoặc tạo ra thực vật hoặc phần của thực vật chứa axit nucleic này, trong đó axit nucleic này chứa vùng điều hòa có hoạt tính ở thực vật và vùng tăng cường sự biểu hiện có hoạt tính ở thực vật, vùng điều hòa và vùng tăng cường sự biểu hiện này liên kết theo cách hoạt động được với trình tự nucleotit mã hóa hemagglutinin (HA) cúm được cải biến typ A typ phụ H3 hoặc H7, HA được cải biến này chứa vòng phân giải protein được xóa bỏ hoàn toàn giữa các dưới đơn vị HA1 và HA2, vòng phân giải protein được xóa bỏ hoàn toàn này chứa vị trí phân cắt monobazơ hoặc đa bazơ, và trong đó axit nucleic này không bao gồm yếu tố khuếch đại virut vàng lùn đậu (BeYDV); và

b) ủ thực vật hoặc phần của thực vật trong các điều kiện mà cho phép sự biểu hiện của axit nucleic, do đó tạo ra VLP.

18. Phương pháp theo điểm 17, trong đó ở bước a) axit nucleic thứ hai được đưa vào trong

thực vật hoặc phần của thực vật, axit nucleic thứ hai này bao gồm vùng điều hoà thứ hai có hoạt tính ở thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá protein kênh proton.

19. Phương pháp theo điểm 18, trong đó protein kênh proton được chọn từ cúm M2 hoặc BM2.

20. Phương pháp theo điểm 17, 18 hoặc 19, còn bao gồm các bước:

c) thu hoạch thực vật, và

d) tinh chế VLP, trong đó VLP có kích cỡ nằm trong khoảng từ 80 đến 300nm.

21. Phương pháp theo điểm bất kỳ trong số các điểm từ 17 đến 20, trong đó ở bước đưa vào (bước a), axit nucleic được đưa vào trong thực vật theo cách tạm thời, hoặc axit nucleic được đưa vào thực vật sao cho nó ổn định.

22. HA được mã hóa bởi axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 16.

23. VLP chứa HA được mã hóa bởi axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 16.

24. VLP theo điểm 23 còn chứa một hoặc nhiều hơn một lipit thu được từ thực vật.

25. VLP theo điểm 24, trong đó VLP này chứa protein HA với N-glycan đặc hiệu thực vật hoặc N-glycan được cải biến.

26. Phương pháp tạo ra protein HA được cải biến bao gồm vòng phân giải protein được xóa bỏ hoàn toàn giữa các dưới đơn vị HA1 và HA2, vòng phân giải protein này bao gồm một hoặc nhiều hơn một vị trí phân cắt monobazơ hoặc đa bazơ thể hiện việc phân cắt giảm hoặc huỷ bỏ ở thực vật, bao gồm các bước:

a) đưa axit nucleic vào trong thực vật, axit nucleic này bao gồm vùng điều hoà có hoạt tính ở thực vật và vùng tăng cường sự biểu hiện có hoạt tính ở thực vật, vùng điều

hoà và vùng tăng cường sự biểu hiện này được liên kết theo cách hoạt động được với trình tự nucleotit mã hoá HA cúm được cải biến typ A typ phụ H3 hoặc H7, HA được cải biến này bao gồm vòng phân giải protein được xóa bỏ hoàn toàn giữa các dưới đơn vị HA1 và HA2, vòng phân giải protein được xóa bỏ hoàn toàn này chứa vị trí phân cắt monobazơ hoặc đa bazơ, và trong đó axit nucleic này không bao gồm yếu tố khuếch đại virut vàng lùn đậu (Bean Yellow Dwarf Virus - BeYDV);

b) ủ thực vật hoặc phần của thực vật trong các điều kiện mà cho phép sự biểu hiện của protein HA, do đó tạo ra protein HA được cải biến, và

c) thu hoạch thực vật và tinh chế protein HA được cải biến.

27. Thực vật chứa axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 16.

28. Chế phẩm chứa liều hữu hiệu của VLP theo điểm 23, 24 hoặc 25 để gây ra đáp ứng miễn dịch và chất mang dược dụng.

29. Vacxin bao gồm liều hữu hiệu của VLP theo điểm 23, 24 hoặc 25 để gây ra đáp ứng miễn dịch.

30. Hemagglutinin (HA) cúm typ B được cải biến hoặc HA cúm typ A typ phụ H3 hoặc H7 chứa vòng phân giải protein được xóa bỏ hoàn toàn giữa các dưới đơn vị HA1 và HA2, vòng phân giải protein được xóa bỏ hoàn toàn này chứa vị trí phân cắt monobazơ hoặc đa bazơ, trong đó HA cúm typ B được cải biến không bao gồm chủng B/Brisbane/60/2008 hoặc chủng B/Wisconsin/1/2010.

31. HA cúm được cải biến theo điểm 30, trong đó HA cúm được cải biến có nguồn gốc từ HA typ B chưa được cải biến chứa 90-100% trình tự giống với trình tự được xác định bởi SEQ ID NO: 37.

32. HA cúm được cải biến theo điểm 30, trong đó HA cúm được cải biến này có nguồn gốc từ HA H3 chưa được cải biến chứa 90-100% trình tự giống với trình tự được xác định bởi SEQ ID NO: 19, các axit amin từ 25 đến 574 của SEQ ID NO: 122 hoặc các axit amin từ 25 đến 274 của SEQ ID NO: 139.

33. HA cúm được cải biến theo điểm 30, trong đó HA cúm được cải biến này có nguồn gốc từ HA H7 chưa được cải biến chứa 90-100% trình tự giống với trình tự được xác định bởi các axit amin từ 25 đến 566 của SEQ ID NO: 150.

34. Chế phẩm chứa lượng hữu hiệu HA được cải biến theo điểm bất kỳ trong số điểm từ 30 đến 33 để tạo ra đáp ứng miễn dịch, và chất mang dược dụng.

35. Vacxin chứa lượng hữu hiệu HA cải biến theo điểm bất kỳ trong số các điểm từ 30 to 33 để tạo ra đáp ứng miễn dịch.

36. Phương pháp làm tăng chất lượng sản phẩm của protein HA được biểu hiện ở thực vật, bao gồm các bước:

a) đưa axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 16 vào trong thực vật;

b) ủ thực vật hoặc phần của thực vật trong các điều kiện mà cho phép sự biểu hiện của protein HA, do đó tạo ra protein HA được cải biến,

c) thu hoạch thực vật và tinh chế protein HA được cải biến, trong đó protein HA được cải biến có chất lượng sản phẩm được gia tăng, khi so với HA nguyên thể.

37. Phương pháp theo điểm 36, trong đó chất lượng bao gồm sản lượng sản phẩm, độ ổn định của sản phẩm, tính bền vững của sản phẩm hoặc tổ hợp của chúng gia tăng.

38. VLP chứa HA cúm được cải biến như được xác định trong điểm bất kỳ trong số các điểm từ 30 đến 33.

DANH MỤC TRÌNH TỰ

<110> MEDICAGO INC.

<120> HẠT TƯƠNG TỰ VIRUT (VLP) CÚM Ở THỰC VẬT, PHƯƠNG PHÁP SẢN XUẤT HẠT TƯƠNG TỰ VIRUT, AXIT NUCLEIC ĐƯỢC SỬ DỤNG TRONG SẢN XUẤT HẠT TƯƠNG TỰ VIRUT VÀ VACXIN VÀ CHẾ PHẨM CHỨA HẠT TƯƠNG TỰ VIRUT

<130> G656-127HNPT

<140> PCT/CA2014/050326
<141> 2014-03-28

<150> US 61/806,227
<151> 2013-03-28

<150> US 61/971,274
<151> 2014-03-27

<150> US 61/925,852
<151> 2014-01-10

<160> 201

<170> PatentIn version 3.5

<210> 1
<211> 21
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự thống nhất vòng phân giải protein H5 ở loài chim

<400> 1

Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile Ala
1 5 10 15

Gly Phe Ile Glu Gly
20

<210> 2
<211> 41
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi IF-H5A-I-05.s1+3c

<400> 2
aaatttgtcg ggcccatgga gaaaatagtg cttcttcttg c 41

<210> 3
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi IF-H5dTm.r

<400> 3
actaaagaaa ataggccttt aaatgcaaat tctgcattgt aacgatccat 50

<210> 4
<211> 4903
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1191

<400> 4
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120

aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180

aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240

ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300

gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360

aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420

aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480

taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga 540

aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600

atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660

taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcatttta 720

tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg 780

gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840

acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat 900

ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960

accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt 1020

gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080

aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg 1140

gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg 1200

actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc 1260

aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg 1320

gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca 1380

tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt 1440

agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg 1500

tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga 1560

tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt 1620

ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa 1680

tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac 1740

ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg 1800

cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa 1860

gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt 1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct 1980

ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc 2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg 2100

cgcgttggga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tggtggagca 2160

cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat 2220

tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gcccagctat 2280

ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat gccatcattg 2340

cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtggtccca aagatggacc 2400

cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt 2460

ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga 2520

tacagtctca gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa 2580

cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga 2640

aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc 2700

tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga 2760

cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga 2820

tgacgcacaa tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca 2880

tttggagagg tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa 2940

ccaaaccttc ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc 3000

ttgcgtgagc gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgttttctt 3060

tcactgaagc gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg 3120

tgtacttgtc ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct 3180

gttcagcccc atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct 3240

acttctgctt gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt 3300

tctataagaa atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga 3360

gaaagattgt taagcttctg tatattctgc ccaaatttgt cgggcccgcg gatggcgaaa 3420

aacgttgcga ttttcggctt attgttttct cttcttgtgt tggttccttc tcagatcttc 3480

gcctgcaggc tcctcagcca aaacgacacc cccatctgtc tatccactgg cccctggatc 3540

tgctgcccaa actaactcca tggtgaccct gggatgcctg gtcaagggct atttccctga 3600

gccagtgaca gtgacctgga actctggatc cctgtccagc ggtgtgcaca ccttcccagc 3660

tgtcctgcag tctgacctct acactctgag cagctcagtg actgtcccct ccagcacctg 3720

gcccagcgag accgtcacct gcaacgttgc ccacccggcc agcagcacca aggtggacaa 3780

gaaaattgtg cccagggatt gtggttgtaa gccttgcata tgtacagtcc cagaagtatc 3840

atctgtcttc atcttccccc caaagcccaa ggatgtgctc accattactc tgactcctaa 3900

ggtcacgtgt gttgtggtag acatcagcaa ggatgatccc gaggtccagt tcagctggtt 3960

tgtagatgat gtggaggtgc acacagctca gacgcaaccc cgggaggagc agttcaacag 4020

cactttccgc tcagtcagtg aacttcccat catgcaccag gactggctca atggcaagga 4080

gcgatcgctc accatcacca tcaccatcac catcaccatt aaaggcctat tttctttagt 4140

ttgaatttac tgttattcgg tgtgcatttc tatgtttggt gagcggtttt ctgtgctcag 4200

agtgtgttta ttttatgtaa tttaatttct ttgtgagctc ctgtttagca ggtcgtccct 4260

tcagcaagga cacaaaaaga ttttaatttt attaaaaaaa aaaaaaaaaa agaccgggaa 4320

ttcgatatca agcttatcga cctgcagatc gttcaaacat ttggcaataa agtttcttaa 4380

gattgaatcc tgttgccggt cttgcgatga ttatcatata atttctgttg aattacgtta 4440

agcatgtaat aattaacatg taatgcatga cgttatttat gagatgggtt tttatgatta 4500

gagtcccgca attatacatt taatacgcga tagaaaacaa aatatagcgc gcaaactagg 4560

ataaattatc gcgcgcggtg tcatctatgt tactagatct ctagagtctc aagcttggcg 4620

cgcccacgtg actagtggca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg 4680

gcgttaccca acttaatcgc cttgcagcac atccccctttt cgccagctgg cgtaatagcg 4740

aagaggcccg caccgatcgc ccttcccaac agttgcgcag cctgaatggc gaatgctaga 4800

gcagcttgag cttggatcag attgtcgttt cccgccttca gtttaaacta tcagtgtttg 4860

acaggatata ttggcgggta aacctaagag aaaagagcgt tta 4903

<210> 5
<211> 3447
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 489

<400> 5
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc agctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatggag aaaatagtgc ttcttcttgc aatagtcagt cttgttaaaa gtgatcagat 1320

ttgcattggt taccatgcaa acaattcaac agagcaggtt gacacaatca tggaaaagaa 1380

cgttactgtt acacatgccc aagacatact ggaaaagaca cacaacggga agctctgcga 1440

tctagatgga gtgaagcctc taattttaag agattgtagt gtagctggat ggctcctcgg 1500

gaacccaatg tgtgacgaat tcatcaatgt accggaatgg tcttacatag tggagaaggc 1560

caatccaacc aatgacctct gttacccagg gagtttcaac gactatgaag aactgaaaca 1620

cctattgagc agaataaacc attttgagaa aattcaaatc atccccaaaa gttcttggtc 1680

cgatcatgaa gcctcatcag gagttagctc agcatgtcca tacctgggaa gtcctcctt 1740

ttttagaaat gtggtatggc ttatcaaaaa gaacagtaca tacccaacaa taaagaaaag 1800

ctacaataat accaaccaag aggatctttt ggtactgtgg ggaattcacc atcctaatga 1860

tgcggcagag cagacaaggc tatatcaaaa cccaaccacc tatatttcca ttgggacatc 1920

aacactaaac cagagattgg taccaaaaat agctactaga tccaaagtaa acgggcaaag 1980

tggaaggatg gagttcttct ggacaatttt aaaacctaat gatgcaatca acttcgagag 2040

taatggaaat ttcattgctc cagaatatgc atacaaaatt gtcaagaaag gggactcagc 2100

aattatgaaa agtgaattgg aatatggtaa ctgcaacacc aagtgtcaaa ctccaatggg 2160

ggcgataaac tctagtatgc cattccacaa catacaccct ctcaccatcg gggaatgccc 2220

caaatatgtg aaatcaaaca gattagtcct tgcaacaggg ctcagaaata gccctcaaag 2280

agagagcaga agaaaaaaga gaggactatt tggagctata gcaggtttta tagagggagg 2340

atggcaggga atggtagatg gttggtatgg gtaccaccat agcaatgagc aggggagtgg 2400

gtacgctgca gacaaagaat ccactcaaaa ggcaatagat ggagtcacca ataaggtcaa 2460

ctcaatcatt gacaaaatga acactcagtt tgaggccgtt ggaagggaat ttaataactt 2520

agaaaggaga atagagaatt taaacaagaa gatggaagac gggtttctag atgtctggac 2580

ttataatgcc gaacttctgg ttctcatgga aaatgagaga actctagact ttcatgactc 2640

aaatgttaag aacctctacg acaaggtccg actacagctt agggataatg caaaggagct 2700

gggtaacggt tgtttcgagt tctatcacaa atgtgataat gaatgtatgg aaagtataag 2760

aaacggaacg tacaactatc cgcagtattc agaagaagca agattaaaaa gagaggaaat 2820

aagtggggta aaattggaat caataggaac ttaccaaata ctgtcaattt attcaacagt 2880

ggcgagttcc ctagcactgg caatcatgat ggctggtcta tctttatgga tgtgctccaa 2940

tggatcgtta caatgcagaa tttgcattta aaggcctatt ttctttagtt tgaatttact 3000

gttattcggt gtgcatttct atgtttggtg agcggttttc tgtgctcaga gtgtgtttat 3060

tttatgtaat ttaatttctt tgtgagctcc tgtttagcag gtcgtccctt cagcaaggac 3120

acaaaaagat tttaatttta ttaaaaaaaa aaaaaaaaaa gaccgggaat tcgatatcaa 3180

gcttatcgac ctgcagatcg ttcaaacatt tggcaataaa gtttcttaag attgaatcct 3240

gttgccggtc ttgcgatgat tatcatataa tttctgttga attacgttaa gcatgtaata 3300

attaacatgt aatgcatgac gttatttatg agatgggttt ttatgattag agtcccgcaa 3360

ttatacattt aatacgcgat agaaaacaaa atatagcgcg caaactagga taaattatcg 3420

cgcgcggtgt catctatgtt actagat 3447

<210> 6
<211> 568
<212> PRT
<213> Cúm H5 A/Indonesia/5/2005 (H5N1)

<400> 6

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
1 5 10 15

```
Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
            20                  25                  30

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
        35                  40                  45

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
    50                  55                  60

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
65                  70                  75                  80

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
                85                  90                  95

Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn
            100                 105                 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
        115                 120                 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
    130                 135                 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
145                 150                 155                 160

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
                165                 170                 175

Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
            180                 185                 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
        195                 200                 205

Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
    210                 215                 220

Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
```

225 230 235 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
245 250 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
260 265 270

Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
275 280 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
290 295 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
305 310 315 320

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
325 330 335

Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
340 345 350

Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
355 360 365

Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
370 375 380

Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
385 390 395 400

Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
405 410 415

Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
420 425 430

Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
435 440 445

Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
450 455 460

Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
465 470 475 480

Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
485 490 495

Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala
500 505 510

Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
515 520 525

Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
530 535 540

Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
545 550 555 560

Ser Leu Gln Cys Arg Ile Cys Ile
565

<210> 7
<211> 42
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi IF-S1-M1+M2ANC.c

<400> 7
aaatttgtcg ggcccatgag tcttctaacc gaggtcgaaa cg 42

<210> 8
<211> 43
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi IF-S1-4-M2ANC.r

<400> 8

actaaagaaa ataggccttt actccagctc tatgctgaca aaa 43

<210> 9
<211> 294
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Gen M2 tổng hợp

<400> 9
atgagtcttc taaccgaggt cgaaacgcct atcagaaacg aatgggggtg cagatgcaac 60

gattcaagtg atcctcttgt tgttgccgca agtataattg ggattgtgca cctgatattg 120

tggattattg atcgcctttt ttccaaaagc atttatcgta tctttaaaca cggtttaaaa 180

agagggcctt ctacggaagg agtaccagag tctatgaggg aagaatatcg agaggaacag 240

cagaatgctg tggatgctga cgatggtcat tttgtcagca tagagctgga gtaa 294

<210> 10
<211> 2034
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 1261

<400> 10
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatgagt cttctaaccg aggtcgaaac gcctatcaga aacgaatggg ggtgcagatg 1320

caacgattca agtgatcctc ttgttgttgc cgcaagtata attgggattg tgcacctgat 1380

attgtggatt attgatcgcc ttttttccaa aagcatttat cgtatcttta aacacggttt 1440

aaaaagaggg ccttctacgg aaggagtacc agagtctatg agggaagaat atcgagagga 1500

acagcagaat gctgtggatg ctgacgatgg tcattttgtc agcatagagc tggagtaaag 1560

gcctattttc tttagtttga atttactgtt attcggtgtg catttctatg tttggtgagc 1620

ggttttctgt gctcagagtg tgtttatttt atgtaattta atttctttgt gagctcctgt 1680

ttagcaggtc gtcccttcag caaggacaca aaaagatttt aattttatta aaaaaaaaaa 1740

aaaaaaagac cgggaattcg atatcaagct tatcgacctg cagatcgttc aaacatttgg 1800

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 1860

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 1920

tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 1980

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agat 2034

<210> 11
<211> 97
<212> PRT
<213> Cúm M2 New Caledonia/20/1999 (H1N1)

<400> 11

Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly
1 5 10 15

Cys Arg Cys Asn Asp Ser Ser Asp Pro Leu Val Val Ala Ala Ser Ile
20 25 30

Ile Gly Ile Val His Leu Ile Leu Trp Ile Ile Asp Arg Leu Phe Ser
35 40 45

Lys Ser Ile Tyr Arg Ile Phe Lys His Gly Leu Lys Arg Gly Pro Ser
50 55 60

Thr Glu Gly Val Pro Glu Ser Met Arg Glu Glu Tyr Arg Glu Glu Gln
65 70 75 80

Gln Asn Ala Val Asp Ala Asp Asp Gly His Phe Val Ser Ile Glu Leu
85 90 95

Glu

<210> 12
<211> 294
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Gen M2 tổng hợp

<400> 12
atgagtcttc taaccgaggt cgaaacgcct atcagaaacg aatgggggtg cagatgcaac 60

ggttcaagtg atcctctcac tattgccgca aatatcattg ggatcttgca cttgacattg 120

tggattcttg atcgtctttt tttcaaatgc atttaccgtc gctttaaata cggactgaaa 180

ggagggcctt ctacggaagg agtgccaaag tctatgaggg aagaatatcg aaaggaacag 240

cagagtgctg tggatgctga cgatggtcat tttgtcagca tagagctgga gtaa 294

<210> 13
<211> 2034
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 859

<400> 13

gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatgagt cttctaaccg aggtcgaaac gcctatcaga aacgaatggg ggtgcagatg 1320

caacggttca agtgatcctc tcactattgc cgcaaatatc attgggatct tgcacttgac 1380

attgtggatt cttgatcgtc ttttttcaa atgcatttac cgtcgcttta aatacggact 1440

gaaaggaggg ccttctacgg aaggagtgcc aaagtctatg agggaagaat atcgaaagga 1500

acagcagagt gctgtggatg ctgacgatgg tcattttgtc agcatagagc tggagtaaag 1560

gcctattttc tttagtttga atttactgtt attcggtgtg catttctatg tttggtgagc 1620

ggttttctgt gctcagagtg tgtttatttt atgtaattta atttctttgt gagctcctgt 1680

ttagcaggtc gtcccttcag caaggacaca aaaagatttt aattttatta aaaaaaaaaa 1740

aaaaaaagac cgggaattcg atatcaagct tatcgacctg cagatcgttc aaacatttgg 1800

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 1860

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 1920

tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 1980

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agat 2034

<210> 14
<211> 97
<212> PRT
<213> Cúm M2 A/Puerto Rico/8/1934 (H1N1)

<400> 14

Met Ser Leu Leu Thr Glu Val Glu Thr Pro Ile Arg Asn Glu Trp Gly
1 5 10 15

Cys Arg Cys Asn Gly Ser Ser Asp Pro Leu Thr Ile Ala Ala Asn Ile
20 25 30

Ile Gly Ile Leu His Leu Thr Leu Trp Ile Leu Asp Arg Leu Phe Phe
35 40 45

Lys Cys Ile Tyr Arg Arg Phe Lys Tyr Gly Leu Lys Gly Gly Pro Ser
50 55 60

Thr Glu Gly Val Pro Lys Ser Met Arg Glu Glu Tyr Arg Lys Glu Gln
65 70 75 80

Gln Ser Ala Val Asp Ala Asp Asp Gly His Phe Val Ser Ile Glu Leu
85 90 95

Glu

<210> 15
<211> 3474
<212> ADN

<213> Trình tự nhân tạo

<220>
<223> Cátxét 1039

<400> 15
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatggcg aaaaacgttg cgattttcgg cttattgttt tctcttcttg tgttggttcc 1320

ttctcagatc ttcgccgatc gaatctgcac tggaataaca tcgtcaaact caccacatgt 1380

cgtcaaaact gctactcaag gggaggtcaa tgtgactggt gtaataccac tgacaacaac 1440

acccaccaaa tctcattttg caaatctcaa aggaacagaa accaggggga aactatgccc 1500

aaaatgcctc aactgcacag atctggacgt agccttgggc agaccaaaat gcacggggaa 1560

aataccctcg gcaagagttt caatactcca tgaagtcaga cctgttacat ctgggtgctt 1620

tcctataatg cacgacagaa caaaaattag acagctgcct aaccttctcc gaggatacga 1680

acatatcagg ttatcaaccc ataacgttat caatgcagaa aatgcaccag gaggacccta 1740

caaaattgga acctcagggt cttgccctaa cattaccaat ggaaacggat ttttcgcaac 1800

aatggcttgg gccgtcccaa aaaacgacaa aaacaaaaca gcaacaaatc cattaacaat 1860

agaagtacca tacatttgta cagaaggaga agaccaaatt accgtttggg ggttccactc 1920

tgacaacgag acccaaatgg caaagctcta tggggactca aagccccaga agttcacctc 1980

atctgccaac ggagtgacca cacattacgt ttcacagatt ggtggcttcc caaatcaaac 2040

agaagacgga ggactaccac aaagtggtag aattgttgtt gattacatgg tgcaaaaatc 2100

tgggaaaaca ggaacaatta cctatcaaag gggtatttta ttgcctcaaa aggtgtggtg 2160

cgcaagtggc aggagcaagg taataaaagg atccttgcct ttaattggag aagcagattg 2220

cctccacgaa aaatacggtg gattaaacaa aagcaagcct tactacacag gggaacatgc 2280

aaaggccata ggaaattgcc caatatgggt gaaaacaccc ttgaagctgg ccaatggaac 2340

caaatataga cctcctggtg gaggatggga aggaatgatt gcaggttggc acggatacac 2400

atcccatggg gcacatggag tagcggtggc agcagacctt aagagcactc aagaggccat 2460

aaacaagata acaaaaaatc tcaactcttt gagtgagctg gaagtaaaga atcttcaaag 2520

actaagcggt gccatggatg aactccacaa cgaaatacta gaactagatg agaaagtgga 2580

tgatctcaga gctgatacaa taagctcaca aatagaactc gcagtcctgc tttccaatga 2640

aggaataata aacagtgaag atgaacatct cttggcgctt gaaagaaagc tgaagaaaat 2700

gctgggcccc tctgctgtag agatagggaa tggatgcttt gaaaccaaac acaagtgcaa 2760

ccagacctgt ctcgacagaa tagctgctgg tacctttgat gcaggagaat tttctctccc 2820

caccttgat tcactgaata ttactgctgc atctttaaat gacgatggat tggataatca 2880

tactatactg ctttactact caactgctgc ctccagtttg gctgtaacac tgatgatagc 2940

tatctttgtt gtttatatgg tctccagaga caatgtttct tgctccatct gtctataaag 3000

gcctattttc tttagtttga atttactgtt attcggtgtg catttctatg tttggtgagc 3060

ggttttctgt gctcagagtg tgtttatttt atgtaattta atttctttgt gagctcctgt 3120

ttagcaggtc gtcccttcag caaggacaca aaaagatttt aattttatta aaaaaaaaaa 3180

aaaaaaagac cgggaattcg atatcaagct tatcgacctg cagatcgttc aaacatttgg 3240

caataaagtt tcttaagatt gaatcctgtt gccggtcttg cgatgattat catataattt 3300

ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 3360

tgggttttta tgattagagt cccgcaatta tacatttaat acgcgataga aaacaaaata 3420

tagcgcgcaa actaggataa attatcgcgc gcggtgtcat ctatgttact agat 3474

<210> 16
<211> 570
<212> PRT
<213> Cúm B/Brisbane/60/2008

<400> 16

Asp Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val
1 5 10 15

Lys Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu
20 25 30

Thr Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Glu
35 40 45

Thr Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp
50 55 60

Val Ala Leu Gly Arg Pro Lys Cys Thr Gly Lys Ile Pro Ser Ala Arg
65 70 75 80

Val Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro
85 90 95

Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg
100 105 110

Gly Tyr Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu
115 120 125

```
Asn Ala Pro Gly Gly Pro Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro
    130                 135                 140

Asn Ile Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val
145                 150                 155                 160

Pro Lys Asn Asp Lys Asn Lys Thr Ala Thr Asn Pro Leu Thr Ile Glu
                165                 170                 175

Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly
            180                 185                 190

Phe His Ser Asp Asn Glu Thr Gln Met Ala Lys Leu Tyr Gly Asp Ser
        195                 200                 205

Lys Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr
    210                 215                 220

Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu
225                 230                 235                 240

Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly
                245                 250                 255

Lys Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys
            260                 265                 270

Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro
        275                 280                 285

Leu Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn
    290                 295                 300

Lys Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn
305                 310                 315                 320

Cys Pro Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys
                325                 330                 335

Tyr Arg Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala
            340                 345                 350
```

```
Ile Ala Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp
        355                 360                 365

His Gly Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp
    370                 375                 380

Leu Lys Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn
385                 390                 395                 400

Ser Leu Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala
                405                 410                 415

Met Asp Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp
            420                 425                 430

Asp Leu Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu
        435                 440                 445

Leu Ser Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala
    450                 455                 460

Leu Glu Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile
465                 470                 475                 480

Gly Asn Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu
                485                 490                 495

Asp Arg Ile Ala Ala Gly Thr Phe Asp Ala Gly Glu Phe Ser Leu Pro
            500                 505                 510

Thr Phe Asp Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly
        515                 520                 525

Leu Asp Asn His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser
    530                 535                 540

Leu Ala Val Thr Leu Met Ile Ala Ile Phe Val Val Tyr Met Val Ser
545                 550                 555                 560
```

Arg Asp Asn Val Ser Cys Ser Ile Cys Leu
565 570

<210> 17
<211> 554
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Vòng phân giải protein delta trình tự axit amin typ B HA (với cầu liên kết GG)

<400> 17

Asp Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val
1 5 10 15

Lys Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu
20 25 30

Thr Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Glu
35 40 45

Thr Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp
50 55 60

Val Ala Leu Gly Arg Pro Lys Cys Thr Gly Lys Ile Pro Ser Ala Arg
65 70 75 80

Val Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro
85 90 95

Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg
100 105 110

Gly Tyr Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu
115 120 125

Asn Ala Pro Gly Gly Pro Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro
130 135 140

Asn Ile Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val
145 150 155 160

```
Pro Lys Asn Asp Lys Asn Lys Thr Ala Thr Asn Pro Leu Thr Ile Glu
                165                 170                 175

Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly
            180                 185                 190

Phe His Ser Asp Asn Glu Thr Gln Met Ala Lys Leu Tyr Gly Asp Ser
        195                 200                 205

Lys Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr
    210                 215                 220

Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu
225                 230                 235                 240

Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly
                245                 250                 255

Lys Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys
            260                 265                 270

Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro
        275                 280                 285

Leu Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn
    290                 295                 300

Lys Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn
305                 310                 315                 320

Cys Pro Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys
                325                 330                 335

Tyr Arg Pro Pro Gly Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His
            340                 345                 350

Gly Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu
        355                 360                 365

Lys Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser
```

370 375 380

Leu Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met
385 390 395 400

Asp Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp
405 410 415

Leu Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu
420 425 430

Ser Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu
435 440 445

Glu Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile Gly
450 455 460

Asn Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp
465 470 475 480

Arg Ile Ala Ala Gly Thr Phe Asp Ala Gly Glu Phe Ser Leu Pro Thr
485 490 495

Phe Asp Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu
500 505 510

Asp Asn His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu
515 520 525

Ala Val Thr Leu Met Ile Ala Ile Phe Val Val Tyr Met Val Ser Arg
530 535 540

Asp Asn Val Ser Cys Ser Ile Cys Leu Pro
545 550

<210> 18
<211> 570
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin thay thế vị trí tách của B HA với cầu liên kết

<400> 18

Asp Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val
1 5 10 15

Lys Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu
20 25 30

Thr Thr Thr Pro Thr Lys Ser His Phe Ala Asn Leu Lys Gly Thr Glu
35 40 45

Thr Arg Gly Lys Leu Cys Pro Lys Cys Leu Asn Cys Thr Asp Leu Asp
50 55 60

Val Ala Leu Gly Arg Pro Lys Cys Thr Gly Lys Ile Pro Ser Ala Arg
65 70 75 80

Val Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro
85 90 95

Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg
100 105 110

Gly Tyr Glu His Ile Arg Leu Ser Thr His Asn Val Ile Asn Ala Glu
115 120 125

Asn Ala Pro Gly Gly Pro Tyr Lys Ile Gly Thr Ser Gly Ser Cys Pro
130 135 140

Asn Ile Thr Asn Gly Asn Gly Phe Phe Ala Thr Met Ala Trp Ala Val
145 150 155 160

Pro Lys Asn Asp Lys Asn Lys Thr Ala Thr Asn Pro Leu Thr Ile Glu
165 170 175

Val Pro Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly
180 185 190

Phe His Ser Asp Asn Glu Thr Gln Met Ala Lys Leu Tyr Gly Asp Ser
195 200 205

```
Lys Pro Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr
    210                 215                 220

Val Ser Gln Ile Gly Gly Phe Pro Asn Gln Thr Glu Asp Gly Gly Leu
225                 230                 235                 240

Pro Gln Ser Gly Arg Ile Val Val Asp Tyr Met Val Gln Lys Ser Gly
                245                 250                 255

Lys Thr Gly Thr Ile Thr Tyr Gln Arg Gly Ile Leu Leu Pro Gln Lys
            260                 265                 270

Val Trp Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro
        275                 280                 285

Leu Ile Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn
    290                 295                 300

Lys Ser Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn
305                 310                 315                 320

Cys Pro Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys
                325                 330                 335

Tyr Arg Gly Ser Ser Ser Gly Ser Ser Ser Gly Gly Phe Phe Gly Ala
            340                 345                 350

Ile Ala Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp
        355                 360                 365

His Gly Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp
    370                 375                 380

Leu Lys Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn
385                 390                 395                 400

Ser Leu Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala
                405                 410                 415

Met Asp Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp
```

420 425 430

Asp Leu Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu
435 440 445

Leu Ser Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala
450 455 460

Leu Glu Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Glu Ile
465 470 475 480

Gly Asn Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu
485 490 495

Asp Arg Ile Ala Ala Gly Thr Phe Asp Ala Gly Glu Phe Ser Leu Pro
500 505 510

Thr Phe Asp Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly
515 520 525

Leu Asp Asn His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser
530 535 540

Leu Ala Val Thr Leu Met Ile Ala Ile Phe Val Val Tyr Met Val Ser
545 550 555 560

Arg Asp Asn Val Ser Cys Ser Ile Cys Leu
565 570

<210> 19
<211> 550
<212> PRT
<213> Cúm H3 A/Perth/16/2009

<400> 19

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
1 5 10 15

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp
20 25 30

```
Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
        35                  40                  45

Gly Glu Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly Lys Asn Cys
    50                  55                  60

Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln
65                  70                  75                  80

Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Tyr Ser Asn
                85                  90                  95

Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
            100                 105                 110

Ala Ser Ser Gly Thr Leu Glu Phe Asn Asn Glu Ser Phe Asn Trp Thr
        115                 120                 125

Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Ile Arg Arg Ser Lys
    130                 135                 140

Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Asn Phe Lys
145                 150                 155                 160

Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Gln Phe Asp Lys
                165                 170                 175

Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Lys Asp Gln Ile
            180                 185                 190

Phe Leu Tyr Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg
        195                 200                 205

Ser Gln Gln Thr Val Ser Pro Asn Ile Gly Ser Arg Pro Arg Val Arg
    210                 215                 220

Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
225                 230                 235                 240

Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly
                245                 250                 255
```

```
Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
            260                 265                 270

Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
        275                 280                 285

Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr Tyr Gly Ala
    290                 295                 300

Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
305                 310                 315                 320

Arg Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala
                325                 330                 335

Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp Gly Trp Tyr Gly
            340                 345                 350

Phe Arg His Gln Asn Ser Glu Gly Arg Gly Gln Ala Ala Asp Leu Lys
        355                 360                 365

Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys Leu Asn Arg Leu
    370                 375                 380

Ile Gly Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser
385                 390                 395                 400

Glu Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr
                405                 410                 415

Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu
            420                 425                 430

Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe
        435                 440                 445

Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn Ala Glu Asp Met Gly Asn
    450                 455                 460
```

Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Gly Ser
465 470 475 480

Ile Arg Asn Gly Thr Tyr Asp His Asp Val Tyr Arg Asp Glu Ala Leu
485 490 495

Asn Asn Arg Phe Gln Ile Lys Gly Val Glu Leu Lys Ser Gly Tyr Lys
500 505 510

Asp Trp Ile Leu Trp Ile Ser Phe Ala Ile Ser Cys Phe Leu Leu Cys
515 520 525

Val Ala Leu Leu Gly Phe Ile Met Trp Ala Cys Gln Lys Gly Asn Ile
530 535 540

Arg Cys Asn Ile Cys Ile
545 550

<210> 20
<211> 540
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Vòng phân giải protein delta trình tự axit amin H3 (với cầu liên kết GS)

<400> 20

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
1 5 10 15

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp
20 25 30

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
35 40 45

Gly Glu Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly Lys Asn Cys
50 55 60

Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln
65 70 75 80

Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Tyr Ser Asn
85 90 95

Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
100 105 110

Ala Ser Ser Gly Thr Leu Glu Phe Asn Asn Glu Ser Phe Asn Trp Thr
115 120 125

Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Ile Arg Arg Ser Lys
130 135 140

Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Asn Phe Lys
145 150 155 160

Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Gln Phe Asp Lys
165 170 175

Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Lys Asp Gln Ile
180 185 190

Phe Leu Tyr Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg
195 200 205

Ser Gln Gln Thr Val Ser Pro Asn Ile Gly Ser Arg Pro Arg Val Arg
210 215 220

Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
225 230 235 240

Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly
245 250 255

Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
260 265 270

Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
275 280 285

Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr Tyr Gly Ala

290 295 300

Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
305 310 315 320

Gly Ser Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu Gly Met
325 330 335

Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly Arg Gly
340 345 350

Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn
355 360 365

Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys Phe His Gln
370 375 380

Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln Asp Leu Glu
385 390 395 400

Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu
405 410 415

Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr Asp Ser
420 425 430

Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn
435 440 445

Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys Cys Asp
450 455 460

Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His Asp Val
465 470 475 480

Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly Val Glu
485 490 495

Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp Ile Ser Phe Ala Ile
500 505 510

Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly Phe Ile Met Trp Ala
515 520 525

Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
530 535 540

<210> 21
<211> 550
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin thay thế vị trí tách của H3 với cầu liên kết

<400> 21

Gln Lys Leu Pro Gly Asn Asp Asn Ser Thr Ala Thr Leu Cys Leu Gly
1 5 10 15

His His Ala Val Pro Asn Gly Thr Ile Val Lys Thr Ile Thr Asn Asp
20 25 30

Gln Ile Glu Val Thr Asn Ala Thr Glu Leu Val Gln Ser Ser Ser Thr
35 40 45

Gly Glu Ile Cys Asp Ser Pro His Gln Ile Leu Asp Gly Lys Asn Cys
50 55 60

Thr Leu Ile Asp Ala Leu Leu Gly Asp Pro Gln Cys Asp Gly Phe Gln
65 70 75 80

Asn Lys Lys Trp Asp Leu Phe Val Glu Arg Ser Lys Ala Tyr Ser Asn
85 90 95

Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Arg Ser Leu Val
100 105 110

Ala Ser Ser Gly Thr Leu Glu Phe Asn Asn Glu Ser Phe Asn Trp Thr
115 120 125

Gly Val Thr Gln Asn Gly Thr Ser Ser Ala Cys Ile Arg Arg Ser Lys
130 135 140

```
Asn Ser Phe Phe Ser Arg Leu Asn Trp Leu Thr His Leu Asn Phe Lys
145                 150                 155                 160

Tyr Pro Ala Leu Asn Val Thr Met Pro Asn Asn Glu Gln Phe Asp Lys
                165                 170                 175

Leu Tyr Ile Trp Gly Val His His Pro Gly Thr Asp Lys Asp Gln Ile
            180                 185                 190

Phe Leu Tyr Ala Gln Ala Ser Gly Arg Ile Thr Val Ser Thr Lys Arg
        195                 200                 205

Ser Gln Gln Thr Val Ser Pro Asn Ile Gly Ser Arg Pro Arg Val Arg
    210                 215                 220

Asn Ile Pro Ser Arg Ile Ser Ile Tyr Trp Thr Ile Val Lys Pro Gly
225                 230                 235                 240

Asp Ile Leu Leu Ile Asn Ser Thr Gly Asn Leu Ile Ala Pro Arg Gly
                245                 250                 255

Tyr Phe Lys Ile Arg Ser Gly Lys Ser Ser Ile Met Arg Ser Asp Ala
            260                 265                 270

Pro Ile Gly Lys Cys Asn Ser Glu Cys Ile Thr Pro Asn Gly Ser Ile
        275                 280                 285

Pro Asn Asp Lys Pro Phe Gln Asn Val Asn Arg Ile Thr Tyr Gly Ala
    290                 295                 300

Cys Pro Arg Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr Gly Met
305                 310                 315                 320

Gly Ser Ser Gly Ser Ser Gly Ser Ser Gly Ile Phe Gly Ala Ile Ala
                325                 330                 335

Gly Phe Ile Glu Asn Gly Trp Glu Gly Met Val Asp Gly Trp Tyr Gly
            340                 345                 350

Phe Arg His Gln Asn Ser Glu Gly Arg Gly Gln Ala Ala Asp Leu Lys
```

355 360 365

Ser Thr Gln Ala Ala Ile Asp Gln Ile Asn Gly Lys Leu Asn Arg Leu
370 375 380

Ile Gly Lys Thr Asn Glu Lys Phe His Gln Ile Glu Lys Glu Phe Ser
385 390 395 400

Glu Val Glu Gly Arg Ile Gln Asp Leu Glu Lys Tyr Val Glu Asp Thr
405 410 415

Lys Ile Asp Leu Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Leu Glu
420 425 430

Asn Gln His Thr Ile Asp Leu Thr Asp Ser Glu Met Asn Lys Leu Phe
435 440 445

Glu Lys Thr Lys Lys Gln Leu Arg Glu Asn Ala Glu Asp Met Gly Asn
450 455 460

Gly Cys Phe Lys Ile Tyr His Lys Cys Asp Asn Ala Cys Ile Gly Ser
465 470 475 480

Ile Arg Asn Gly Thr Tyr Asp His Asp Val Tyr Arg Asp Glu Ala Leu
485 490 495

Asn Asn Arg Phe Gln Ile Lys Gly Val Glu Leu Lys Ser Gly Tyr Lys
500 505 510

Asp Trp Ile Leu Trp Ile Ser Phe Ala Ile Ser Cys Phe Leu Leu Cys
515 520 525

Val Ala Leu Leu Gly Phe Ile Met Trp Ala Cys Gln Lys Gly Asn Ile
530 535 540

Arg Cys Asn Ile Cys Ile
545 550

<210> 22
<211> 31
<212> PRT

<213> Cúm H1 New Cal

<400> 22

Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
1 5 10 15

Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr
20 25 30

<210> 23
<211> 31
<212> PRT
<213> Cúm H1 Brisbane

<400> 23

Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
1 5 10 15

Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr
20 25 30

<210> 24
<211> 31
<212> PRT
<213> Cúm H1 Sol Islands

<400> 24

Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
1 5 10 15

Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr
20 25 30

<210> 25
<211> 31
<212> PRT
<213> Cúm H2A Singapore

<400> 25

Val Pro Gln Ile Glu Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
1 5 10 15

Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr

20 25 30

<210> 26
<211> 31
<212> PRT
<213> Cúm H3A Brisbane

<400> 26

Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala Gly Phe
1 5 10 15

Ile Glu Asn Gly Trp Glu Gly Met Val Asp Gly Trp Tyr Gly Phe
20 25 30

<210> 27
<211> 31
<212> PRT
<213> Cúm H3A WCN

<400> 27

Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala Gly Phe
1 5 10 15

Ile Glu Asn Gly Trp Glu Gly Met Val Asp Gly Trp Tyr Gly Phe
20 25 30

<210> 28
<211> 34
<212> PRT
<213> Cúm H5 Anhui

<400> 28

Ser Pro Leu Arg Glu Arg Arg Arg Lys Arg Gly Leu Phe Gly Ala Ile
1 5 10 15

Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
20 25 30

Gly Tyr

<210> 29
<211> 35

<212> PRT
<213> Cúm H5 Indo

<400> 29

Ser Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala
1 5 10 15

Ile Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp
20 25 30

Tyr Gly Tyr
35

<210> 30
<211> 35
<212> PRT
<213> Cúm H5 Vietnam

<400> 30

Ser Pro Gln Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala
1 5 10 15

Ile Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp
20 25 30

Tyr Gly Tyr
35

<210> 31
<211> 6861
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1194

<400> 31
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120

aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180

aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240

ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300

gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360

aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420

aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480

taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga 540

aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600

atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660

taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcattttta 720

tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg 780

gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840

acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat 900

ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960

accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt 1020

gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080

aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg 1140

gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg 1200

actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc 1260

aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg 1320

gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca 1380

tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt 1440

agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg 1500

tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga 1560

tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt 1620

ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa 1680

tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac 1740

ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg 1800

cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa 1860

gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt 1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct 1980

ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc 2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg 2100

cgcgttggga attactagcg cgtgtcgaca cgcgtggcgc gccctagcag aaggcatgtt 2160

gttgtgactc cgaggggttg cctcaaactc tatcttataa ccggcgtgga ggcatggagg 2220

caagggcatt ttggtaattt aagtagttag tggaaaatga cgtcatttac ttaaagacga 2280

agtcttgcga caaggggggc ccacgccgaa ttttaatatt accggcgtgg ccccacctta 2340

tcgcgagtgc tttagcacga gcggtccaga tttaaagtag aaaagttccc gcccactagg 2400

gttaaaggtg ttcacactat aaaagcatat acgatgtgat ggtatttgat aaagcgtata 2460

ttgtatcagg tatttccgtc ggatacgaat tattcgtaca agcttcttaa gccggtcaac 2520

atggtggagc acgacacact tgtctactcc aaaaatatca aagatacagt ctcagaagac 2580

caaagggcaa ttgagacttt tcaacaaagg gtaatatccg gaaacctcct cggattccat 2640

tgcccagcta tctgtcactt tattgtgaag atagtggaaa aggaaggtgg ctcctacaaa 2700

tgccatcatt gcgataaagg aaaggccatc gttgaagatg cctctgccga cagtggtccc 2760

aaagatggac ccccacccac gaggagcatc gtggaaaaag aagacgttcc aaccacgtct 2820

tcaaagcaag tggattgatg tgataacatg gtggagcacg acacacttgt ctactccaaa 2880

aatatcaaag atacagtctc agaagaccaa agggcaattg agacttttca acaaagggta 2940

atatccggaa acctcctcgg attccattgc ccagctatct gtcactttat tgtgaagata 3000

gtggaaaagg aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggccatcgtt 3060

gaagatgcct ctgccgacag tggtcccaaa gatggacccc cacccacgag gagcatcgtg 3120

gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga tatctccact 3180

gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc tatataagga 3240

agttcatttc atttggagag gtattaaaat cttaataggt tttgataaaa gcgaacgtgg 3300

ggaaacccga accaaacctt cttctaaact ctctctcatc tctcttaaag caaacttctc 3360

tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga tcgtgcttcg gcaccagtac 3420

aacgttttct ttcactgaag cgaaatcaaa gatctctttg tggacacgta gtgcggcgcc 3480

attaaataac gtgtacttgt cctattcttg tcggtgtggt cttgggaaaa gaaagcttgc 3540

tggaggctgc tgttcagccc catacattac ttgttacgat tctgctgact ttcggcgggt 3600

gcaatatctc tacttctgct tgacgaggta ttgttgcctg tacttctttc ttcttcttct 3660

tgctgattgg ttctataaga aatctagtat tttctttgaa acagagtttt cccgtggttt 3720

tcgaacttgg agaaagattg ttaagcttct gtatattctg cccaaatttg tcgggcccat 3780

ggcgaaaaac gttgcgattt tcggcttatt gttttctctt cttgtgttgg ttccttctca 3840

gatcttcgcc gcggctcctc agccaaaacg acacccccat ctgtctatcc actggcccct 3900

ggatctgctg cccaaactaa ctccatggtg accctgggat gcctggtcaa gggctatttc 3960

cctgagccag tgacagtgac ctggaactct ggatccctgt ccagcggtgt gcacaccttc 4020

ccagctgtcc tgcagtctga cctctacact ctgagcagct cagtgactgt cccctccagc 4080

acctggccca gcgagaccgt cacctgcaac gttgcccacc cggccagcag caccaaggtg 4140

gacaagaaaa ttgtgcccag ggattgtggt tgtaagcctt gcatatgtac agtcccagaa 4200

gtatcatctg tcttcatctt ccccccaaag cccaaggatg tgctcaccat tactctgact 4260

cctaaggtca cgtgtgttgt ggtagacatc agcaaggatg atcccgaggt ccagttcagc 4320

tggtttgtag atgatgtgga ggtgcacaca gctcagacgc aaccccggga ggagcagttc 4380

aacagcactt tccgctcagt cagtgaactt cccatcatgc accaggactg gctcaatggc 4440

aaggaaggcc tattttcttt agtttgaatt tactgttatt cggtgtgcat ttctatgttt 4500

ggtgagcggt tttctgtgct cagagtgtgt ttattttatg taatttaatt tctttgtgag 4560

ctcctgttta gcaggtcgtc ccttcagcaa ggacacaaaa agattttaat tttattaaaa 4620

aaaaaaaaaa aaaagaccgg gaattcgata tcaagcttat cgacctgcag atcgttcaaa 4680

catttggcaa taaagtttct taagattgaa tcctgttgcc ggtcttgcga tgattatcat 4740

ataatttctg ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt 4800

tatgagatgg gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa 4860

caaaatatag cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga 4920

tctctagagt ctcaagcttg gcgcggggta ccgagctcga attccgagtg tacttcaagt 4980

cagttggaaa tcaataaaat gattatttta tgaatatatt tcattgtgca agtagataga 5040

aattacatat gttacataac acacgaaata aacaaaaaaa cacaatccaa aacaaacacc 5100

ccaaacaaaa taacactata tatatcctcg tatgaggaga ggcacgttca gtgactcgac 5160

gattcccgag caaaaaaagt ctccccgtca cacatatagt gggtgacgca attatcttca 5220

aagtaatcct tctgttgact tgtcattgat aacatccagt cttcgtcagg attgcaaaga 5280

attatagaag ggatcccacc ttttattttc ttctttttc catatttagg gttgacagtg 5340

aaatcagact ggcaacctat taattgcttc cacaatggga cgaacttgaa ggggatgtcg 5400

tcgatgatat tataggtggc gtgttcatcg tagttggtga agtcgatggt cccgttccag 5460

tagttgtgtc gcccgagact tctagcccag gtggtctttc cggtacgagt tggtccgcag 5520

atgtagaggc tggggtgtct gaccccagtc cttccctcat cctggttaga tcggccatcc 5580

actcaaggtc agattgtgct tgatcgtagg agacaggatg tatgaaagtg taggcatcga 5640

tgcttacatg atataggtgc gtctctctcc agttgtgcag atcttcgtgg cagcggagat 5700

ctgattctgt gaagggcgac acgtactgct caggttgtgg aggaaataat ttgttggctg 5760

aatattccag ccattgaagc tttgttgccc attcatgagg gaattcttct ttgatcatgt 5820

caagatactc ctccttagac gttgcagtct ggataatagt tcgccatcgt gcgtcagatt 5880

tgcgaggaga gaccttatga tctcggaaat ctcctctggt tttaatatct ccgtcctttg 5940

atatgtaatc aaggacttgt ttagagtttc tagctggctg gatattaggg tgatttcctt 6000

caaaatcgaa aaaagaagga tccctaatac aaggtttttt atcaagctgg ataagagcat 6060

gatagtgggt agtgccatct tgatgaagct cagaagcaac accaaggaag aaaataagaa 6120

aaggtgtgag tttctcccag agaaactgga ataaatcatc tctttgagat gagcacttgg 6180

ggtaggtaag gaaaacatat ttagattgga gtctgaagtt cttgctagca gaaggcatgt 6240

tgttgtgact ccgaggggtt gcctcaaact ctatcttata accggcgtgg aggcatggag 6300

gcaagggcat tttggtaatt taagtagtta gtggaaaatg acgtcattta cttaaagacg 6360

aagtcttgcg acaaggggggg cccacgccga attttaatat taccggcgtg gccccacctt 6420

atcgcgagtg ctttagcacg agcggtccag atttaaagta gaaaagttcc cgcccactag 6480

ggttaaaggt gttcacacta taaaagcata tacgatgtga tggtatttga tggagcgtat 6540

attgtatcag gtatttccgt cggatacgaa ttattcgtac ggccggccac tagtggcact 6600

ggccgtcgtt ttacaacgtc gtgactggga aaaccctggc gttacccaac ttaatcgcct 6660

tgcagcacat ccccctttcg ccagctggcg taatagcgaa gaggcccgca ccgatcgccc 6720

ttcccaacag ttgcgcagcc tgaatggcga atgctagagc agcttgagct tggatcagat 6780

tgtcgtttcc cgccttcagt ttaaactatc agtgtttgac aggatatatt ggcgggtaaa 6840

cctaagagaa aagagcgttt a 6861

<210> 32
<211> 5555
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 1008

<400> 32
ctagcagaag gcatgttgtt gtgactccga ggggttgcct caaactctat cttataaccg 60

gcgtggaggc atggaggcaa gggcattttg gtaatttaag tagttagtgg aaaatgacgt 120

catttactta aagacgaagt cttgcgacaa ggggggccca cgccgaattt taatattacc 180

ggcgtggccc caccttatcg cgagtgcttt agcacgagcg gtccagattt aaagtagaaa 240

agttcccgcc cactagggtt aaaggtgttc acactataaa agcatatacg atgtgatggt 300

atttgataaa gcgtatattg tatcaggtat ttccgtcgga tacgaattat tcgtacaagc 360

ttcttaagcc ggtcaacatg gtggagcacg acacacttgt ctactccaaa aatatcaaag 420

atacagtctc agaagaccaa agggcaattg agacttttca acaaagggta atatccggaa 480

acctcctcgg attccattgc ccagctatct gtcactttat tgtgaagata gtggaaaagg 540

aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggccatcgtt gaagatgcct 600

ctgccgacag tggtcccaaa gatggacccc cacccacgag gagcatcgtg gaaaaagaag 660

acgttccaac cacgtcttca aagcaagtgg attgatgtga taacatggtg gagcacgaca 720

cacttgtcta ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga 780

cttttcaaca aagggtaata tccggaaacc tcctcggatt ccattgccca gctatctgtc 840

actttattgt gaagatagtg gaaaaggaag gtggctccta caaatgccat cattgcgata 900

aaggaaaggc catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac 960

ccacgaggag catcgtggaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt 1020

gatgtgatat ctccactgac gtaagggatg acgcacaatc ccactatcct tcgcaagacc 1080

cttcctctat ataaggaagt tcatttcatt tggagaggta ttaaaatctt aataggtttt 1140

gataaaagcg aacgtgggga aacccgaacc aaaccttctt ctaaactctc tctcatctct 1200

cttaaagcaa acttctctct tgtctttctt gcgtgagcga tcttcaacgt tgtcagatcg 1260

tgcttcggca ccagtacaac gttttctttc actgaagcga aatcaaagat ctctttgtgg 1320

acacgtagtg cggcgccatt aaataacgtg tacttgtcct attcttgtcg gtgtggtctt 1380

gggaaaagaa agcttgctgg aggctgctgt tcagccccat acattacttg ttacgattct 1440

gctgactttc ggcgggtgca atatctctac ttctgcttga cgaggtattg ttgcctgtac 1500

ttctttcttc ttcttcttgc tgattggttc tataagaaat ctagtatttt ctttgaaaca 1560

gagttttccc gtggttttcg aacttggaga aagattgtta agcttctgta tattctgccc 1620

aaatttgtcg ggcccatggc gaaaaacgtt gcgattttcg gcttattgtt ttctcttctt 1680

gtgttggttc cttctcagat cttcgccgat cgaatctgca ctggaataac atcgtcaaac 1740

tcaccacatg tcgtcaaaac tgctactcaa ggggaggtca atgtgactgg tgtaatacca 1800

ctgacaacaa cacccaccaa atctcatttt gcaaatctca aaggaacaga aaccaggggg 1860

aaactatgcc caaaatgcct caactgcaca gatctggacg tagccttggg cagaccaaaa 1920

tgcacgggga aaataccctc ggcaagagtt tcaatactcc atgaagtcag acctgttaca 1980

tctgggtgct ttcctataat gcacgacaga acaaaaatta gacagctgcc taaccttctc 2040

cgaggatacg aacatatcag gttatcaacc cataacgtta tcaatgcaga aaatgcacca 2100

ggaggaccct acaaaattgg aacctcaggg tcttgcccta acattaccaa tggaaacgga 2160

tttttcgcaa caatggcttg ggccgtccca aaaaacgaca aaaacaaaac agcaacaaat 2220

ccattaacaa tagaagtacc atacatttgt acagaaggag aagaccaaat taccgtttgg 2280

gggttccact ctgacaacga gacccaaatg gcaaagctct atggggactc aaagccccag 2340

aagttcacct catctgccaa cggagtgacc acacattacg tttcacagat tggtggcttc 2400

ccaaatcaaa cagaagacgg aggactacca caaagtggta gaattgttgt tgattacatg 2460

gtgcaaaaat ctgggaaaac aggaacaatt acctatcaaa ggggtatttt attgcctcaa 2520

aaggtgtggt gcgcaagtgg caggagcaag gtaataaaag gatccttgcc tttaattgga 2580

gaagcagatt gcctccacga aaaatacggt ggattaaaca aaagcaagcc ttactacaca 2640

ggggaacatg caaaggccat aggaaattgc ccaatatggg tgaaaacacc cttgaagctg 2700

gccaatggaa ccaaatatag acctcctgca aaactattaa aggaaagggg tttcttcgga 2760

gctattgctg gtttcttaga aggaggatgg gaaggaatga ttgcaggttg gcacggatac 2820

acatcccatg gggcacatgg agtagcggtg gcagcagacc ttaagagcac tcaagaggcc 2880

ataaacaaga taacaaaaaa tctcaactct ttgagtgagc tggaagtaaa gaatcttcaa 2940

agactaagcg gtgccatgga tgaactccac aacgaaatac tagaactaga tgagaaagtg 3000

gatgatctca gagctgatac aataagctca caaatagaac tcgcagtcct gctttccaat 3060

gaaggaataa taaacagtga agatgaacat ctcttggcgc ttgaaagaaa gctgaagaaa 3120

atgctgggcc cctctgctgt agagataggg aatggatgct ttgaaaccaa acacaagtgc 3180

aaccagacct gtctcgacag aatagctgct ggtacctttg atgcaggaga attttctctc 3240

cccacctttg attcactgaa tattactgct gcatctttaa atgacgatgg attggataat 3300

catactatac tgctttacta ctcaactgct gcctccagtt tggctgtaac actgatgata 3360

gctatctttg ttgtttatat ggtctccaga gacaatgttt cttgctccat ctgtctataa 3420

aggcctattt tctttagttt gaatttactg ttattcggtg tgcatttcta tgtttggtga 3480

gcggttttct gtgctcagag tgtgtttatt ttatgtaatt taatttcttt gtgagctcct 3540

gtttagcagg tcgtcccttc agcaaggaca caaaaagatt ttaattttat taaaaaaaaa 3600

aaaaaaaaag accgggaatt cgatatcaag cttatcgacc tgcagatcgt tcaaacattt 3660

ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt atcatataat 3720

ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg ttatttatga 3780

gatgggtttt tatgattaga gtcccgcaat tatacattta atacgcgata gaaaacaaaa 3840

tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta ctagatctct 3900

agagtctcaa gcttggcgcg gggtaccgag ctcgaattcc gagtgtactt caagtcagtt 3960

ggaaatcaat aaaatgatta ttttatgaat atatttcatt gtgcaagtag atagaaatta 4020

catatgttac ataacacacg aaataaacaa aaaaacacaa tccaaaacaa acaccccaaa 4080

caaaataaca ctatatatat cctcgtatga ggagaggcac gttcagtgac tcgacgattc 4140

ccgagcaaaa aaagtctccc cgtcacacat atagtgggtg acgcaattat cttcaaagta 4200

atccttctgt tgacttgtca ttgataacat ccagtcttcg tcaggattgc aaagaattat 4260

agaagggatc ccacctttta ttttcttctt ttttccatat ttagggttga cagtgaaatc 4320

agactggcaa cctattaatt gcttccacaa tgggacgaac ttgaagggga tgtcgtcgat 4380

gatattatag gtggcgtgtt catcgtagtt ggtgaagtcg atggtcccgt tccagtagtt 4440

gtgtcgcccg agacttctag cccaggtggt ctttccggta cgagttggtc cgcagatgta 4500

gaggctgggg tgtctgaccc cagtccttcc ctcatcctgg ttagatcggc catccactca 4560

aggtcagatt gtgcttgatc gtaggagaca ggatgtatga aagtgtaggc atcgatgctt 4620

acatgatata ggtgcgtctc tctccagttg tgcagatctt cgtggcagcg gagatctgat 4680

tctgtgaagg gcgacacgta ctgctcaggt tgtggaggaa ataatttgtt ggctgaatat 4740

tccagccatt gaagctttgt tgcccattca tgagggaatt cttctttgat catgtcaaga 4800

tactcctcct tagacgttgc agtctggata atagttcgcc atcgtgcgtc agatttgcga 4860

ggagagacct tatgatctcg gaaatctcct ctggttttaa tatctccgtc ctttgatatg 4920

taatcaagga cttgtttaga gtttctagct ggctggatat tagggtgatt tccttcaaaa 4980

tcgaaaaaag aaggatccct aatacaaggt tttttatcaa gctggataag agcatgatag 5040

tgggtagtgc catcttgatg aagctcagaa gcaacaccaa ggaagaaaat aagaaaaggt 5100

gtgagtttct cccagagaaa ctggaataaa tcatctcttt gagatgagca cttggggtag 5160

gtaaggaaaa catatttaga ttggagtctg aagttcttgc tagcagaagg catgttgttg 5220

tgactccgag gggttgcctc aaactctatc ttataaccgg cgtggaggca tggaggcaag 5280

ggcattttgg taatttaagt agttagtgga aaatgacgtc atttacttaa agacgaagtc 5340

ttgcgacaag gggggcccac gccgaatttt aatattaccg gcgtggcccc accttatcgc 5400

gagtgcttta gcacgagcgg tccagattta aagtagaaaa gttcccgccc actagggtta 5460

aaggtgttca cactataaaa gcatatacga tgtgatggta tttgatggag cgtatattgt 5520

atcaggtatt tccgtcggat acgaattatt cgtac 5555

<210> 33
<211> 31
<212> PRT
<213> Cúm H6 Teal HK

<400> 33

Val Pro Gln Ile Glu Thr Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
1 5 10 15

Ile Glu Gly Gly Trp Thr Gly Met Ile Asp Gly Trp Tyr Gly Tyr
20 25 30

<210> 34
<211> 42
<212> PRT
<213> Cúm H7 Eq Prague

<400> 34

Val Pro Glu Ala Pro Ala His Lys Gln Leu Thr His His Met Arg Lys
1 5 10 15

Lys Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp
20 25 30

Glu Gly Leu Ile Asp Gly Trp Tyr Gly Tyr
35 40

<210> 35
<211> 31
<212> PRT
<213> Cúm H9A HK

<400> 35

Val Pro Ala Arg Ser Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
1 5 10 15

Ile Glu Gly Gly Trp Pro Gly Leu Val Ala Gly Trp Tyr Gly Phe
20 25 30

<210> 36
<211> 43
<212> PRT
<213> Cúm B Florida

<400> 36

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys Leu Leu
1 5 10 15

Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu Gly Gly
20 25 30

Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr
35 40

<210> 37
<211> 43
<212> PRT
<213> Cúm B Malaysia

<400> 37

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys Leu Leu
1 5 10 15

Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu Gly Gly
20 25 30

Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr
35 40

<210> 38
<211> 52
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1059 1039 cộng với 1059.r

<400> 38
cttcccatcc tccaccagga ggtctatatt tggttccatt ggccagcttc aa 52

<210> 39
<211> 52
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1039 cộng với 1059.c

<400> 39
caaatataga cctcctggtg gaggatggga aggaatgatt gcaggttggc ac 52

<210> 40
<211> 5504
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 1059

<400> 40
ctagcagaag gcatgttgtt gtgactccga ggggttgcct caaactctat cttataaccg 60

gcgtggaggc atggaggcaa gggcattttg gtaatttaag tagttagtgg aaaatgacgt 120

catttactta aagacgaagt cttgcgacaa ggggggccca cgccgaattt taatattacc 180

ggcgtggccc caccttatcg cgagtgcttt agcacgagcg gtccagattt aaagtagaaa 240

agttcccgcc cactagggtt aaaggtgttc acactataaa agcatatacg atgtgatggt 300

atttgataaa gcgtatattg tatcaggtat ttccgtcgga tacgaattat tcgtacaagc 360

ttcttaagcc ggtcaacatg gtggagcacg acacacttgt ctactccaaa aatatcaaag 420

atacagtctc agaagaccaa agggcaattg agacttttca acaaagggta atatccggaa 480

acctcctcgg attccattgc ccagctatct gtcactttat tgtgaagata gtggaaaagg 540

aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggccatcgtt gaagatgcct 600

ctgccgacag tggtcccaaa gatggacccc cacccacgag gagcatcgtg gaaaaagaag 660

acgttccaac cacgtcttca aagcaagtgg attgatgtga taacatggtg gagcacgaca 720

cacttgtcta ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga 780

cttttcaaca aagggtaata tccggaaacc tcctcggatt ccattgccca gctatctgtc 840

actttattgt gaagatagtg gaaaaggaag gtggctccta caaatgccat cattgcgata 900

aaggaaaggc catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac 960

ccacgaggag catcgtggaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt 1020

gatgtgatat ctccactgac gtaagggatg acgcacaatc ccactatcct tcgcaagacc 1080

cttcctctat ataaggaagt tcatttcatt tggagaggta ttaaaatctt aataggtttt 1140

gataaaagcg aacgtgggga aacccgaacc aaaccttctt ctaaactctc tctcatctct 1200

cttaaagcaa acttctctct tgtctttctt gcgtgagcga tcttcaacgt tgtcagatcg 1260

tgcttcggca ccagtacaac gttttctttc actgaagcga aatcaaagat ctctttgtgg 1320

acacgtagtg cggcgccatt aaataacgtg tacttgtcct attcttgtcg gtgtggtctt 1380

gggaaaagaa agcttgctgg aggctgctgt tcagccccat acattacttg ttacgattct 1440

gctgactttc ggcgggtgca atatctctac ttctgcttga cgaggtattg ttgcctgtac 1500

ttctttcttc ttcttcttgc tgattggttc tataagaaat ctagtatttt ctttgaaaca 1560

gagttttccc gtggttttcg aacttggaga aagattgtta agcttctgta tattctgccc 1620

aaatttgtcg ggcccatggc gaaaaacgtt gcgattttcg gcttattgtt ttctcttctt 1680

gtgttggttc cttctcagat cttcgccgat cgaatctgca ctggaataac atcgtcaaac 1740

tcaccacatg tcgtcaaaac tgctactcaa ggggaggtca atgtgactgg tgtaatacca 1800

ctgacaacaa cacccaccaa atctcatttt gcaaatctca aaggaacaga aaccaggggg 1860

aaactatgcc caaaatgcct caactgcaca gatctggacg tagccttggg cagaccaaaa 1920

tgcacgggga aaataccctc ggcaagagtt tcaatactcc atgaagtcag acctgttaca 1980

tctgggtgct ttcctataat gcacgacaga acaaaaatta gacagctgcc taaccttctc 2040

cgaggatacg aacatatcag gttatcaacc cataacgtta tcaatgcaga aaatgcacca 2100

ggaggaccct acaaaattgg aacctcaggg tcttgcccta acattaccaa tggaaacgga 2160

tttttcgcaa caatggcttg ggccgtccca aaaaacgaca aaaacaaaac agcaacaaat 2220

ccattaacaa tagaagtacc atacatttgt acagaaggag aagaccaaat taccgtttgg 2280

gggttccact ctgacaacga gacccaaatg gcaaagctct atggggactc aaagccccag 2340

aagttcacct catctgccaa cggagtgacc acacattacg tttcacagat tggtggcttc 2400

ccaaatcaaa cagaagacgg aggactacca caaagtggta gaattgttgt tgattacatg 2460

gtgcaaaaat ctgggaaaac aggaacaatt acctatcaaa ggggtatttt attgcctcaa 2520

aaggtgtggt gcgcaagtgg caggagcaag gtaataaaag gatccttgcc tttaattgga 2580

gaagcagatt gcctccacga aaaatacggt ggattaaaca aaagcaagcc ttactacaca 2640

ggggaacatg caaaggccat aggaaattgc ccaatatggg tgaaaacacc cttgaagctg 2700

gccaatggaa ccaaatatag acctcctggt ggaggatggg aaggaatgat tgcaggttgg 2760

cacggataca catcccatgg ggcacatgga gtagcggtgg cagcagacct taagagcact 2820

caagaggcca taaacaagat aacaaaaaat ctcaactctt tgagtgagct ggaagtaaag 2880

aatcttcaaa gactaagcgg tgccatggat gaactccaca acgaaatact agaactagat 2940

gagaaagtgg atgatctcag agctgataca ataagctcac aaatagaact cgcagtcctg 3000

ctttccaatg aaggaataat aaacagtgaa gatgaacatc tcttggcgct tgaaagaaag 3060

ctgaagaaaa tgctgggccc ctctgctgta gagataggga atggatgctt tgaaaccaaa 3120

cacaagtgca accagacctg tctcgacaga atagctgctg gtacctttga tgcaggagaa 3180

ttttctctcc ccacctttga ttcactgaat attactgctg catctttaaa tgacgatgga 3240

ttggataatc atactatact gctttactac tcaactgctg cctccagttt ggctgtaaca 3300

ctgatgatag ctatctttgt tgtttatatg gtctccagag acaatgtttc ttgctccatc 3360

tgtctataaa ggcctatttt ctttagtttg aatttactgt tattcggtgt gcatttctat 3420

gtttggtgag cggttttctg tgctcagagt gtgtttattt tatgtaattt aatttctttg 3480

tgagctcctg tttagcaggt cgtcccttca gcaaggacac aaaaagattt taattttatt 3540

aaaaaaaaaa aaaaaaaaga ccgggaattc gatatcaagc ttatcgacct gcagatcgtt 3600

caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt gcgatgatta 3660

tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa tgcatgacgt 3720

tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa tacgcgatag 3780

aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca tctatgttac 3840

tagatctcta gagtctcaag cttggcgcgg ggtaccgagc tcgaattccg agtgtacttc 3900

aagtcagttg gaaatcaata aaatgattat tttatgaata tatttcattg tgcaagtaga 3960

tagaaattac atatgttaca taacacacga aataaacaaa aaaacacaat ccaaaacaaa 4020

caccccaaac aaaataacac tatatatatc ctcgtatgag gagaggcacg ttcagtgact 4080

cgacgattcc cgagcaaaaa aagtctcccc gtcacacata tagtgggtga cgcaattatc 4140

ttcaaagtaa tccttctgtt gacttgtcat tgataacatc cagtcttcgt caggattgca 4200

aagaattata gaagggatcc cacctttttat tttcttcttt tttccatatt tagggttgac 4260

agtgaaatca gactggcaac ctattaattg cttccacaat gggacgaact tgaaggggat 4320

gtcgtcgatg atattatagg tggcgtgttc atcgtagttg gtgaagtcga tggtcccgtt 4380

ccagtagttg tgtcgcccga gacttctagc ccaggtggtc tttccggtac gagttggtcc 4440

gcagatgtag aggctggggt gtctgacccc agtccttccc tcatcctggt tagatcggcc 4500

atccactcaa ggtcagattg tgcttgatcg taggagacag gatgtatgaa agtgtaggca 4560

tcgatgctta catgatatag gtgcgtctct ctccagttgt gcagatcttc gtggcagcgg 4620

agatctgatt ctgtgaaggg cgacacgtac tgctcaggtt gtggaggaaa taatttgttg 4680

gctgaatatt ccagccattg aagctttgtt gcccattcat gagggaattc ttctttgatc 4740

atgtcaagat actcctcctt agacgttgca gtctggataa tagttcgcca tcgtgcgtca 4800

gatttgcgag gagagacctt atgatctcgg aaatctcctc tggttttaat atctccgtcc 4860

tttgatatgt aatcaaggac ttgtttagag tttctagctg gctggatatt agggtgattt 4920

ccttcaaaat cgaaaaaaga aggatcccta atacaaggtt ttttatcaag ctggataaga 4980

gcatgatagt gggtagtgcc atcttgatga agctcagaag caacaccaag gaagaaaata 5040

agaaaaggtg tgagtttctc ccagagaaac tggaataaat catctctttg agatgagcac 5100

ttggggtagg taaggaaaac atatttagat tggagtctga agttcttgct agcagaaggc 5160

atgttgttgt gactccgagg ggttgcctca aactctatct tataaccggc gtggaggcat 5220

ggaggcaagg gcattttggt aatttaagta gttagtggaa aatgacgtca tttacttaaa 5280

gacgaagtct tgcgacaagg ggggcccacg ccgaatttta atattaccgg cgtggcccca 5340

ccttatcgcg agtgctttag cacgagcggt ccagatttaa agtagaaaag ttcccgccca 5400

ctagggttaa aggtgttcac actataaaag catatacgat gtgatggtat ttgatggagc 5460

gtatattgta tcaggtattt ccgtcggata cgaattattc gtac 5504

<210> 41
<211> 577
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> trình tự axit amin PDISP/HA cúm B/Brisbane/60/2008
(vòng phân giải protein loại bớt)

<400> 41

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
50 55 60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
85 90 95

```
Thr Gly Lys Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
            100                 105                 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
        115                 120                 125

Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
    130                 135                 140

Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
145                 150                 155                 160

Ile Gly Thr Ser Gly Ser Cys Pro Asn Ile Thr Asn Gly Asn Gly Phe
                165                 170                 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Lys Asn Lys Thr
            180                 185                 190

Ala Thr Asn Pro Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly
        195                 200                 205

Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu Thr Gln
    210                 215                 220

Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser
225                 230                 235                 240

Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro
                245                 250                 255

Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val
            260                 265                 270

Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln
        275                 280                 285

Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser
    290                 295                 300
```

```
Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu
305                 310                 315                 320

His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly
                325                 330                 335

Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro
            340                 345                 350

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp
        355                 360                 365

Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His
    370                 375                 380

Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn
385                 390                 395                 400

Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn
                405                 410                 415

Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu
            420                 425                 430

Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser
        435                 440                 445

Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser
    450                 455                 460

Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu
465                 470                 475                 480

Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys His
                485                 490                 495

Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asp
            500                 505                 510

Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala
        515                 520                 525
```

Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu Tyr
530 535 540

Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Ile Ala Ile
545 550 555 560

Phe Val Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile Cys
565 570 575

Leu

<210> 42
<211> 1707
<212> ADN
<213> Cúm H5 A/Indonesia/5/2005 (H5N1)

<400> 42
atggagaaaa tagtgcttct tcttgcaata gtcagtcttg ttaaaagtga tcagatttgc 60

attggttacc atgcaaacaa ttcaacagag caggttgaca caatcatgga aaagaacgtt 120

actgttacac atgcccaaga catactggaa aagacacaca acgggaagct ctgcgatcta 180

gatggagtga agcctctaat tttaagagat tgtagtgtag ctggatggct cctcgggaac 240

ccaatgtgtg acgaattcat caatgtaccg gaatggtctt acatagtgga gaaggccaat 300

ccaaccaatg acctctgtta cccagggagt ttcaacgact atgaagaact gaaacaccta 360

ttgagcagaa taaaccattt tgagaaaatt caaatcatcc ccaaaagttc ttggtccgat 420

catgaagcct catcaggagt tagctcagca tgtccatacc tgggaagtcc ctccttttt 480

agaaatgtgg tatggcttat caaaaagaac agtacatacc caacaataaa gaaaagctac 540

aataatacca accaagagga tcttttggta ctgtggggaa ttcaccatcc taatgatgcg 600

gcagagcaga caaggctata tcaaaaccca accacctata tttccattgg gacatcaaca 660

ctaaaccaga gattggtacc aaaaatagct actagatcca aagtaaacgg gcaaagtgga 720

aggatggagt tcttctggac aattttaaaa cctaatgatg caatcaactt cgagagtaat 780

ggaaatttca ttgctccaga atatgcatac aaaattgtca agaaagggga ctcagcaatt 840

atgaaaagtg aattggaata tggtaactgc aacaccaagt gtcaaactcc aatgggggcg 900

ataaactcta gtatgccatt ccacaacata caccctctca ccatcgggga atgccccaaa 960

tatgtgaaat caaacagatt agtccttgca acagggctca gaaatagccc tcaaagagag 1020

agcagaagaa aaaagagagg actatttgga gctatagcag gttttataga gggaggatgg 1080

cagggaatgg tagatggttg gtatgggtac caccatagca atgagcaggg gagtgggtac 1140

gctgcagaca aagaatccac tcaaaaggca atagatggag tcaccaataa ggtcaactca 1200

atcattgaca aaatgaacac tcagtttgag gccgttggaa gggaatttaa taacttagaa 1260

aggagaatag agaatttaaa caagaagatg gaagacgggt ttctagatgt ctggacttat 1320

aatgccgaac ttctggttct catggaaaat gagagaactc tagactttca tgactcaaat 1380

gttaagaacc tctacgacaa ggtccgacta cagcttaggg ataatgcaaa ggagctgggt 1440

aacggttgtt tcgagttcta tcacaaatgt gataatgaat gtatggaaag tataagaaac 1500

ggaacgtaca actatccgca gtattcagaa gaagcaagat taaaaagaga ggaaataagt 1560

ggggtaaaat tggaatcaat aggaacttac caaatactgt caatttattc aacagtggcg 1620

agttccctag cactggcaat catgatggct ggtctatctt tatggatgtg ctccaatgga 1680

tcgttacaat gcagaatttg catttaa 1707

<210> 43
<211> 1734
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> trình tự nucleotit PDISP/HA cúm B/Brisbane/60/2008
(vòng phân giải protein loại bớt)

<400> 43
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg ccgatcgaat ctgcactgga ataacatcgt caaactcacc acatgtcgtc 120

aaaactgcta ctcaagggga ggtcaatgtg actggtgtaa taccactgac aacaacaccc 180

accaaatctc attttgcaaa tctcaaagga acagaaacca gggggaaact atgcccaaaa 240

tgcctcaact gcacagatct ggacgtagcc ttgggcagac caaaatgcac ggggaaaata 300

ccctcggcaa gagtttcaat actccatgaa gtcagacctg ttacatctgg gtgctttcct 360

ataatgcacg acagaacaaa aattagacag ctgcctaacc ttctccgagg atacgaacat 420

atcaggttat caacccataa cgttatcaat gcagaaaatg caccaggagg accctacaaa 480

attggaacct cagggtcttg ccctaacatt accaatggaa acggattttt cgcaacaatg 540

gcttgggccg tcccaaaaaa cgacaaaaac aaaacagcaa caaatccatt aacaatagaa 600

gtaccataca tttgtacaga aggagaagac caaattaccg tttgggggtt ccactctgac 660

aacgagaccc aaatggcaaa gctctatggg gactcaaagc cccagaagtt cacctcatct 720

gccaacggag tgaccacaca ttacgtttca cagattggtg gcttcccaaa tcaaacagaa 780

gacggaggac taccacaaag tggtagaatt gttgttgatt acatggtgca aaaatctggg 840

aaaacaggaa caattaccta tcaaaggggt attttattgc ctcaaaaggt gtggtgcgca 900

agtggcagga gcaaggtaat aaaaggatcc ttgcctttaa ttggagaagc agattgcctc 960

cacgaaaaat acggtggatt aaacaaaagc aagccttact acacagggga acatgcaaag 1020

gccataggaa attgcccaat atgggtgaaa acacccttga agctggccaa tggaaccaaa 1080

tatagacctc ctggtggagg atgggaagga atgattgcag gttggcacgg atacacatcc 1140

catggggcac atggagtagc ggtggcagca gaccttaaga gcactcaaga ggccataaac 1200

aagataacaa aaaatctcaa ctctttgagt gagctggaag taaagaatct tcaaagacta 1260

agcggtgcca tggatgaact ccacaacgaa atactagaac tagatgagaa agtggatgat 1320

ctcagagctg atacaataag ctcacaaata gaactcgcag tcctgctttc caatgaagga 1380

ataataaaca gtgaagatga acatctcttg gcgcttgaaa gaaagctgaa gaaaatgctg 1440

ggcccctctg ctgtagagat agggaatgga tgctttgaaa ccaaacacaa gtgcaaccag 1500

acctgtctcg acagaatagc tgctggtacc tttgatgcag gagaattttc tctccccacc 1560

tttgattcac tgaatattac tgctgcatct ttaaatgacg atggattgga taatcatact 1620

atactgcttt actactcaac tgctgcctcc agtttggctg taacactgat gatagctatc 1680

tttgttgttt atatggtctc cagagacaat gtttcttgct ccatctgtct ataa 1734

<210> 44
<211> 19
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> trình tự tự nhiên của vị trí tách H5/Indo

<400> 44

Thr Gly Leu Arg Asn Ser Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg
1 5 10 15

Gly Leu Phe

<210> 45
<211> 15
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách cải biến H5/Indo (TETR)

<400> 45

Thr Gly Leu Arg Asn Ser Pro Gln Thr Glu Thr Arg Gly Leu Phe
1 5 10 15

<210> 46
<211> 15
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách cải biến H5/Indo (TETQ)

<400> 46

Thr Gly Leu Arg Asn Ser Pro Gln Thr Glu Thr Gln Gly Leu Phe
1 5 10 15

<210> 47
<211> 11
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách H1

<400> 47

Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe
1 5 10

<210> 48
<211> 11
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách H3

<400> 48

Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe
1 5 10

<210> 49
<211> 46
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1462 IF-HAB110.S1+3c

<400> 49

Ala Ala Ala Thr Thr Thr Gly Thr Cys Gly Gly Gly Cys Cys Cys Ala
1 5 10 15

Thr Gly Ala Ala Gly Gly Cys Ala Ala Thr Ala Ala Thr Thr Gly Thr
20 25 30

Ala Cys Thr Ala Cys Thr Cys Ala Thr Gly Gly Thr Ala Gly
35 40 45

<210> 50
<211> 52
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc số 1462 IF-HAB110.s1-4r

<400> 50
actaaagaaa ataggccttt atagacagat ggagcatgaa acgttgtctc tg 52

<210> 51
<211> 1755
<212> ADN
<213> Cúm HA B wisconsin

<400> 51
atgaaggcaa taattgtact actcatggta gtaacatcca atgcagatcg aatctgcact 60

gggataacat cttcaaactc acctcatgtg gtcaaaacag ctactcaagg ggaggtcaat 120

gtgactggcg tgataccact gacaacaaca ccaacaaaat cttattttgc aaatctcaaa 180

ggaacaagga ccagagggaa actatgcccg gactgtctca actgtacaga tctggatgtg 240

gccttgggca ggccaatgtg tgtggggacc acaccttctg ctaaagcttc aatactccac 300

gaggtcagac ctgttacatc cgggtgcttt cctataatgc acgacagaac aaaaatcagg 360

caactaccca atcttctcag aggatatgaa aatatcaggt tatcaaccca aaacgttatc 420

gatgcagaaa aagcaccagg aggaccctac agacttggaa cctcaggatc ttgccctaac 480

gctaccagta aaatcggatt ttttgcaaca atggcttggg ctgtcccaaa ggacaactac 540

aaaaatgcaa cgaacccact aacagtagaa gtaccataca tttgtacaga aggggaagac 600

caaattactg tttgggggtt ccattcagat aacaaaaccc aaatgaagag cctctatgga 660

gactcaaatc ctcaaaagtt cacctcatct gctaatggag taaccacaca ttatgtttct 720

cagattggcg acttcccaga tcaaacagaa gacggaggac taccacaaag cggcagaatt 780

gttgttgatt acatgatgca aaaacctggg aaaacaggaa caattgtcta tcaaagaggt 840

gttttgttgc ctcaaaaggt gtggtgcgcg agtggcagga gcaaagtaat aaaagggtca 900

ttgcctttaa ttggtgaagc agattgcctt catgaaaaat acggtggatt aaacaaaagc 960

aagccttact acacaggaga acatgcaaaa gccataggaa attgcccaat atgggtaaaa 1020

acacctttga agcttgccaa tggaaccaaa tatagacctc ctgcaaaact attgaaggaa 1080

aggggtttct tcggagctat tgctggtttc ctagaaggag gatgggaagg aatgattgca 1140

ggttggcacg gatacacatc tcacggagca catggagtgg cagtggcggc agaccttaag 1200

agtacacaag aagctataaa taagataaca aaaaatctca attctttgag tgagctagaa 1260

gtaaagaacc ttcaaagact aagtggtgcc atggatgaac tccacaacga aatactcgag 1320

ctggatgaga aagtggatga tctcagagct gacactataa gctcacaaat agaacttgca 1380

gtcttgcttt ccaacgaagg aataataaac agtgaagacg agcatctatt ggcacttgag 1440

agaaaactaa agaaaatgct gggtccctct gctgtagaca taggaaacgg atgcttcgaa 1500

accaaacaca aatgcaacca gacctgctta gacaggatag ctgctggcac ctttaatgca 1560

ggagaatttt ctctccccac ttttgattca ttgaacatta ctgctgcatc tttaaatgat 1620

gatggattgg ataaccatac tatactgctc tattactcaa ctgctgcttc tagtttggct 1680

gtaacattaa tgctagctat ttttattgtt tatatggtct ccagagacaa cgtttcatgc 1740

tccatctgtc tataa 1755

<210> 52
<211> 6745
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 193

<400> 52
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120

aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180

aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240

ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300

gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360

aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420

aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480

taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga 540

aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600

atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660

taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcattttta 720

tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg 780

gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840

acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat 900

ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960

accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt 1020

gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080

aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg 1140

gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg 1200

actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc 1260

aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg 1320

gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca 1380

tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt 1440

agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg 1500

tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga 1560

tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt 1620

ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa 1680

tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac 1740

ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg 1800

cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa 1860

gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt 1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct 1980

ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc 2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg 2100

cgcgttggga attactagcg cgtgtcgaga cgcgttgttg ttgtgactcc gaggggttgc 2160

ctcaaactct atcttataac cggcgtggag gcatggaggc aggggtattt tggtcatttt 2220

aatagatagt ggaaaatgac gtggaattta cttaaagacg aagtctttgc gacaaggggg 2280

ggcccacgcc gaatttaata ttaccggcgt ggccccccct tatcgcgagt gctttagcac 2340

gagcggtcca gatttaaagt agaaaatttc ccgcccacta gggttaaagg tgttcacact 2400

ataaaagcat atacgatgtg atggtatttg gtcgacaagc ttgcatgccg gtcaacatgg 2460

tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca gaagaccaaa 2520

gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga ttccattgcc 2580

cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc tacaaatgcc 2640

atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt ggtcccaaag 2700

atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc acgtcttcaa 2760

agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac tccaaaaata 2820

tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa agggtaatat 2880

ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg aagatagtgg 2940

aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc atcgttgaag 3000

atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc atcgtggaaa 3060

aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc tccactgacg 3120

taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata taaggaagtt 3180

catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga acgtggggaa 3240

acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa cttctctctt 3300

gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac cagtacaacg 3360

ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc ggcgccatta 3420

aataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa gcttgctgga 3480

ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg gcgggtgcaa 3540

tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct tcttcttgct 3600

gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg tggttttcga 3660

acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg gcccgcggat 3720

ggcgaaaaac gttgcgattt tcggcttatt gttttctctt cttgtgttgg ttccttctca 3780

gatcttcgcc tgcaggctcc tcagccaaaa cgacaccccc atctgtctat ccactggccc 3840

ctggatctgc tgcccaaact aactccatgg tgaccctggg atgcctggtc aagggctatt 3900

tccctgagcc agtgacagtg acctggaact ctggatccct gtccagcggt gtgcacacct 3960

tcccagctgt cctgcagtct gacctctaca ctctgagcag ctcagtgact gtcccctcca 4020

gcacctggcc cagcgagacc gtcacctgca acgttgccca cccggccagc agcaccaagg 4080

tggacaagaa aattgtgccc agggattgtg gttgtaagcc ttgcatatgt acagtcccag 4140

aagtatcatc tgtcttcatc ttcccccaa agcccaagga tgtgctcacc attactctga 4200

ctcctaaggt cacgtgtgtt gtggtagaca tcagcaagga tgatcccgag gtccagttca 4260

gctggtttgt agatgatgtg gaggtgcaca cagctcagac gcaaccccgg gaggagcagt 4320

tcaacagcac tttccgctca gtcagtgaac ttcccatcat gcaccaggac tggctcaatg 4380

gcaaggagcg atcgctcacc atcaccatca ccatcaccat caccattaaa ggcctatttt 4440

ctttagtttg aatttactgt tattcggtgt gcatttctat gtttggtgag cggttttctg 4500

tgctcagagt gtgtttattt tatgtaattt aatttctttg tgagctcctg tttagcaggt 4560

cgtcccttca gcaaggacac aaaaagattt taattttatt aaaaaaaaaa aaaaaaaaga 4620

ccgggaattc gatatcaagc ttatcgacct gcagatcgtt caaacatttg gcaataaagt 4680

ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 4740

tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggttttt 4800

atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 4860

aactaggata aattatcgcg cgcggtgtca tctatgttac tagatctcta gagtctcaag 4920

cttggcgcgc cataaaatga ttattttatg aatatatttc attgtgcaag tagatagaaa 4980

ttacatatgt tacataacac acgaaataaa caaaaaaaga caatccaaaa acaaacaccc 5040

caaaaaaaat aatcacttta gataaactcg tatgaggaga ggcacgttca gtgactcgac 5100

gattcccgag caaaaaaagt ctccccgtca cacatatagt gggtgacgca attatcttta 5160

aagtaatcct tctgttgact tgtcattgat aacatccagt cttcgtcagg attgcaaaga 5220

attatagaag ggatcccacc ttttattttc ttctttttc catatttagg gttgacagtg 5280

aaatcagact ggcaacctat taattgcttc cacaatggga cgaacttgaa ggggatgtcg 5340

tcgatgatat tataggtggc gtgttcatcg tagttggtga aatcgatggt accgttccaa 5400

tagttgtgtc gtccgagact tctagcccag gtggtctttc cggtacgagt tggtccgcag 5460

atgtagaggc tggggtgtcg gattccattc cttccattgt cctggttaaa tcggccatcc 5520

attcaaggtc agattgagct tgttggtatg agacaggatg tatgtaagta taagcgtcta 5580

tgcttacatg gtatagatgg gtttccctcc aggagtgtag atcttcgtgg cagcgaagat 5640

ctgattctgt gaagggcgac acatacggtt caggttgtgg agggaataat ttgttggctg 5700

aatattccag ccattgaagt tttgttgccc attcatgagg gaattcttcc ttgatcatgt 5760

caagatattc ctccttagac gttgcagtct ggataatagt tctccatcgt gcgtcagatt 5820

tgcgaggaga gaccttatga tctcggaaat ctcctctggt tttaatatct ccgtcctttg 5880

atatgtaatc aaggacttgt ttagagtttc tagctggctg gatattaggg tgatttcctt 5940

caaaatcgaa aaaagaagga tccctaatac aaggtttttt atcaagctgg agaagagcat 6000

gatagtgggt agtgccatct tgatgaagct cagaagcaac accaaggaag aaaataagaa 6060

aaggtgtgag tttctcccag agaaactgga ataaatcatc tctttgagat gagcacttgg 6120

gataggtaag gaaaacatat ttagattgga gtctgaagtt cttactagca gaaggcatgt 6180

tgttgtgact ccgaggggtt gcctcaaact ctatcttata accggcgtgg aggcatggag 6240

gcaggggtat tttggtcatt ttaatagata gtggaaaatg acgtggaatt tacttaaaga 6300

cgaagtcttt gcgacaaggg ggggcccacg ccgaatttaa tattaccggc gtggcccccc 6360

cttatcgcga gtgctttagc acgagcggtc cagatttaaa gtagaaaatt tcccgcccac 6420

tagggttaaa ggtgttcaca ctataaaagc atatacgatg tgatggtatt tgactagtgg 6480

cactggccgt cgttttacaa cgtcgtgact gggaaaaccc tggcgttacc caacttaatc 6540

gccttgcagc acatcccccc ttcgccagct ggcgtaatag cgaagaggcc cgcaccgatc 6600

gcccttccca acagttgcgc agcctgaatg gcgaatgcta gagcagcttg agcttggatc 6660

agattgtcgt ttcccgcctt cagtttaaac tatcagtgtt tgacaggata tattggcggg 6720

taaacctaag agaaaagagc gttta 6745

<210> 53
<211> 3495
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 1462

<400> 53
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatgaag gcaataattg tactactcat ggtagtaaca tccaatgcag atcgaatctg 1320

cactgggata acatcttcaa actcacctca tgtggtcaaa acagctactc aaggggaggt 1380

caatgtgact ggcgtgatac cactgacaac aacaccaaca aaatcttatt ttgcaaatct 1440

caaaggaaca aggaccagag ggaaactatg cccggactgt ctcaactgta cagatctgga 1500

tgtggccttg ggcaggccaa tgtgtgtggg gaccacacct tctgctaaag cttcaatact 1560

ccacgaggtc agacctgtta catccgggtg ctttcctata atgcacgaca gaacaaaaat 1620

caggcaacta cccaatcttc tcagaggata tgaaaatatc aggttatcaa cccaaaacgt 1680

tatcgatgca gaaaaagcac caggaggacc ctacagactt ggaacctcag gatcttgccc 1740

taacgctacc agtaaaatcg gatttttgc aacaatggct tgggctgtcc caaaggacaa 1800

ctacaaaaat gcaacgaacc cactaacagt agaagtacca tacatttgta cagaagggga 1860

agaccaaatt actgtttggg ggttccattc agataacaaa acccaaatga agagcctcta 1920

tggagactca aatcctcaaa agttcacctc atctgctaat ggagtaacca cacattatgt 1980

ttctcagatt ggcgacttcc cagatcaaac agaagacgga ggactaccac aaagcggcag 2040

aattgttgtt gattacatga tgcaaaaacc tgggaaaaca ggaacaattg tctatcaaag 2100

aggtgttttg ttgcctcaaa aggtgtggtg cgcgagtggc aggagcaaag taataaaagg 2160

gtcattgcct ttaattggtg aagcagattg ccttcatgaa aaatacggtg gattaaacaa 2220

aagcaagcct tactacacag gagaacatgc aaaagccata ggaaattgcc caatatgggt 2280

aaaaacacct ttgaagcttg ccaatggaac caaatataga cctcctgcaa aactattgaa 2340

ggaaaggggt ttcttcggag ctattgctgg tttcctagaa ggaggatggg aaggaatgat 2400

tgcaggttgg cacggataca catctcacgg agcacatgga gtggcagtgg cggcagacct 2460

taagagtaca caagaagcta taaataagat aacaaaaaat ctcaattctt tgagtgagct 2520

agaagtaaag aaccttcaaa gactaagtgg tgccatggat gaactccaca acgaaatact 2580

cgagctggat gagaaagtgg atgatctcag agctgacact ataagctcac aaatagaact 2640

tgcagtcttg ctttccaacg aaggaataat aaacagtgaa gacgagcatc tattggcact 2700

tgagagaaaa ctaaagaaaa tgctgggtcc ctctgctgta gacataggaa acggatgctt 2760

cgaaaccaaa cacaaatgca accagacctg cttagacagg atagctgctg gcacctttaa 2820

tgcaggagaa ttttctctcc ccacttttga ttcattgaac attactgctg catctttaaa 2880

tgatgatgga ttggataacc atactatact gctctattac tcaactgctg cttctagttt 2940

ggctgtaaca ttaatgctag ctatttttat tgtttatatg gtctccagag acaacgtttc 3000

atgctccatc tgtctataaa ggcctatttt ctttagtttg aatttactgt tattcggtgt 3060

gcatttctat gtttggtgag cggttttctg tgctcagagt gtgtttattt tatgtaattt 3120

aatttctttg tgagctcctg tttagcaggt cgtcccttca gcaaggacac aaaaagattt 3180

taattttatt aaaaaaaaaa aaaaaaaaga ccgggaattc gatatcaagc ttatcgacct 3240

gcagatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt 3300

gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa 3360

tgcatgacgt tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa 3420

tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca 3480

tctatgttac tagat 3495

<210> 54
<211> 584
<212> PRT
<213> Cúm B/Wisconsin/1/2010

<400> 54

Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
1 5 10 15

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys

20 25 30

Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
35 40 45

Thr Thr Pro Thr Lys Ser Tyr Phe Ala Asn Leu Lys Gly Thr Arg Thr
50 55 60

Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp Val
65 70 75 80

Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys Ala
85 90 95

Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile
100 105 110

Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly
115 120 125

Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu Lys
130 135 140

Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn
145 150 155 160

Ala Thr Ser Lys Ile Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro
165 170 175

Lys Asp Asn Tyr Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val Pro
180 185 190

Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His
195 200 205

Ser Asp Asn Lys Thr Gln Met Lys Ser Leu Tyr Gly Asp Ser Asn Pro
210 215 220

Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
225 230 235 240

```
Gln Ile Gly Asp Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln
                245                 250                 255

Ser Gly Arg Ile Val Val Asp Tyr Met Met Gln Lys Pro Gly Lys Thr
            260                 265                 270

Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp
        275                 280                 285

Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile
    290                 295                 300

Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
305                 310                 315                 320

Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
                325                 330                 335

Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
            340                 345                 350

Pro Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala
        355                 360                 365

Gly Phe Leu Glu Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly
    370                 375                 380

Tyr Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys
385                 390                 395                 400

Ser Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu
                405                 410                 415

Ser Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp
            420                 425                 430

Glu Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu
        435                 440                 445

Arg Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser
```

```
    450                 455                 460

Asn Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu
465                 470                 475                 480

Arg Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Asp Ile Gly Asn
                485                 490                 495

Gly Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg
            500                 505                 510

Ile Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe
        515                 520                 525

Asp Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp
    530                 535                 540

Asn His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala
545                 550                 555                 560

Val Thr Leu Met Leu Ala Ile Phe Ile Val Tyr Met Val Ser Arg Asp
                565                 570                 575

Asn Val Ser Cys Ser Ile Cys Leu
            580
```

<210> 55
<211> 56
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1467 HAB110(PrL-).r

<400> 55
tccttcccat cctccaccag gaggtctata tttggttcca ttggcaagct tcaaag 56

<210> 56
<211> 52
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1467 HAB110(PrL-).c

<400> 56
atatagacct cctggtggag gatgggaagg aatgattgca ggttggcacg ga 52

<210> 57
<211> 3444
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1467

<400> 57
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta ataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatgaag gcaataattg tactactcat ggtagtaaca tccaatgcag atcgaatctg 1320

cactgggata acatcttcaa actcacctca tgtggtcaaa acagctactc aaggggaggt 1380

caatgtgact ggcgtgatac cactgacaac aacaccaaca aaatcttatt ttgcaaatct 1440

caaaggaaca aggaccagag ggaaactatg cccggactgt ctcaactgta cagatctgga 1500

tgtggccttg ggcaggccaa tgtgtgtggg gaccacacct tctgctaaag cttcaatact 1560

ccacgaggtc agacctgtta catccgggtg ctttcctata atgcacgaca gaacaaaaat 1620

caggcaacta cccaatcttc tcagaggata tgaaaatatc aggttatcaa cccaaaacgt 1680

tatcgatgca gaaaaagcac caggaggacc ctacagactt ggaacctcag gatcttgccc 1740

taacgctacc agtaaaatcg gatttttgc aacaatggct tgggctgtcc caaaggacaa 1800

ctacaaaaat gcaacgaacc cactaacagt agaagtacca tacatttgta cagaagggga 1860

agaccaaatt actgtttggg ggttccattc agataacaaa acccaaatga agagcctcta 1920

tggagactca aatcctcaaa agttcacctc atctgctaat ggagtaacca cacattatgt 1980

ttctcagatt ggcgacttcc cagatcaaac agaagacgga ggactaccac aaagcggcag 2040

aattgttgtt gattacatga tgcaaaaacc tgggaaaaca ggaacaattg tctatcaaag 2100

aggtgttttg ttgcctcaaa aggtgtggtg cgcgagtggc aggagcaaag taataaaagg 2160

gtcattgcct ttaattggtg aagcagattg ccttcatgaa aaatacggtg gattaaacaa 2220

aagcaagcct tactacacag gagaacatgc aaaagccata ggaaattgcc caatatgggt 2280

aaaaacacct ttgaagcttg ccaatggaac caaatataga cctcctggtg gaggatggga 2340

aggaatgatt gcaggttggc acggatacac atctcacgga gcacatggag tggcagtggc 2400

ggcagacctt aagagtacac aagaagctat aaataagata acaaaaaatc tcaattcttt 2460

gagtgagcta gaagtaaaga accttcaaag actaagtggt gccatggatg aactccacaa 2520

cgaaatactc gagctggatg agaaagtgga tgatctcaga gctgacacta taagctcaca 2580

aatagaactt gcagtcttgc tttccaacga aggaataata aacagtgaag acgagcatct 2640

attggcactt gagagaaaac taaagaaaat gctgggtccc tctgctgtag acataggaaa 2700

cggatgcttc gaaaccaaac acaaatgcaa ccagacctgc ttagacagga tagctgctgg 2760

caccttaat gcaggagaat tttctctccc cacttttgat tcattgaaca ttactgctgc 2820

atctttaaat gatgatggat tggataacca tactatactg ctctattact caactgctgc 2880

ttctagtttg gctgtaacat taatgctagc tatttttatt gtttatatgg tctccagaga 2940

caacgtttca tgctccatct gtctataaag gcctattttc tttagtttga atttactgtt 3000

attcggtgtg catttctatg tttggtgagc ggttttctgt gctcagagtg tgtttatttt 3060

atgtaattta atttctttgt gagctcctgt ttagcaggtc gtcccttcag caaggacaca 3120

aaaagatttt aattttatta aaaaaaaaaa aaaaaaagac cgggaattcg atatcaagct 3180

tatcgacctg cagatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt 3240

gccggtcttg cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt 3300

aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta 3360

tacatttaat acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc 3420

gcggtgtcat ctatgttact agat 3444

<210> 58
<211> 567
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> trình tự axit amin HA cúm B/Wisconsin/1/2010 (PL loại bớt)

<400> 58

Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
1 5 10 15

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
20 25 30

Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
35 40 45

Thr Thr Pro Thr Lys Ser Tyr Phe Ala Asn Leu Lys Gly Thr Arg Thr
50 55 60

Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp Val
65 70 75 80

Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys Ala

85 90 95

Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile
100 105 110

Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly
115 120 125

Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu Lys
130 135 140

Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn
145 150 155 160

Ala Thr Ser Lys Ile Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro
165 170 175

Lys Asp Asn Tyr Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val Pro
180 185 190

Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His
195 200 205

Ser Asp Asn Lys Thr Gln Met Lys Ser Leu Tyr Gly Asp Ser Asn Pro
210 215 220

Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
225 230 235 240

Gln Ile Gly Asp Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln
245 250 255

Ser Gly Arg Ile Val Val Asp Tyr Met Met Gln Lys Pro Gly Lys Thr
260 265 270

Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp
275 280 285

Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile
290 295 300

```
Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
305                 310                 315                 320

Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
                325                 330                 335

Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
            340                 345                 350

Pro Pro Gly Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr
        355                 360                 365

Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys Ser
    370                 375                 380

Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser
385                 390                 395                 400

Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu
                405                 410                 415

Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg
            420                 425                 430

Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn
        435                 440                 445

Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg
    450                 455                 460

Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Asp Ile Gly Asn Gly
465                 470                 475                 480

Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile
                485                 490                 495

Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp
            500                 505                 510

Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn
```

515 520 525

His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val
530 535 540

Thr Leu Met Leu Ala Ile Phe Ile Val Tyr Met Val Ser Arg Asp Asn
545 550 555 560

Val Ser Cys Ser Ile Cys Leu
565

<210> 59
<211> 11
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách B

<400> 59

Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe
1 5 10

<210> 60
<211> 19
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách tự nhiên H5/Indo

<400> 60

Thr Gly Leu Arg Asn Ser Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg
1 5 10 15

Gly Leu Phe

<210> 61
<211> 15
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách cải biến H5/Indo

<400> 61

Thr Gly Leu Arg Asn Ser Pro Gln Thr Glu Thr Arg Gly Leu Phe
1 5 10 15

<210> 62
<211> 15
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Vị trí tách cải biến H5/Indo

<400> 62

Thr Gly Leu Arg Asn Ser Pro Gln Thr Glu Thr Gln Gly Leu Phe
1 5 10 15

<210> 63
<211> 11
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách cải biến H1/Brisbane

<400> 63

Asn Ile Pro Ser Ile Gln Ser Gln Gly Leu Phe
1 5 10

<210> 64
<211> 11
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách cải biến H3/Brisbane

<400> 64

Asn Val Pro Glu Ile Gln Thr Gln Gly Ile Phe
1 5 10

<210> 65
<211> 11
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> vị trí tách cải biến B/Florida, B/Brisbane

<400> 65

Pro Ala Ile Leu Leu Asn Ile Gln Gly Phe Phe
1 5 10

<210> 66
<211> 19
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> thống nhất A/H3/HA0

<400> 66

Asn Val Pro Glu Lys Gln Thr Arg Gly Ile Phe Gly Ala Ile Ala Gly
1 5 10 15

Phe Ile Glu

<210> 67
<211> 19
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> thống nhất A/H1/HA0

<400> 67

Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
1 5 10 15

Phe Ile Glu

<210> 68
<211> 19
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> thống nhất B/HA0

<400> 68

Pro Ala Lys Leu Leu Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly
1 5 10 15

Phe Leu Glu

<210> 69
<211> 16
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> loại bớt vòng phân giải protein H5 Anhui

<400> 69

Ser Pro Leu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr
1 5 10 15

<210> 70
<211> 16
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> loại bớt vòng phân giải protein H5 Indo

<400> 70

Ser Pro Gln Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr
1 5 10 15

<210> 71
<211> 16
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> loại bớt vòng phân giải protein H5 Vietnam

<400> 71

Ser Pro Gln Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr
1 5 10 15

<210> 72
<211> 25
<212> PRT

<213> Trình tự nhân tạo

<220>
<223> loại bớt vòng phân giải protein B Florida

<400> 72

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Trp Glu
1 5 10 15

Gly Met Ile Ala Gly Trp His Gly Tyr
20 25

<210> 73
<211> 25
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> loại bớt vòng phân giải protein B Malaysia

<400> 73

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Trp Glu
1 5 10 15

Gly Met Ile Ala Gly Trp His Gly Tyr
20 25

<210> 74
<211> 45
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi MutCleavage-H5(Indo).r

<400> 74
tagtcctctt gtctctgttt gagggctatt tctgagccct gttgc 45

<210> 75
<211> 45
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi MutCleavage-H5(Indo).c

<400> 75

tagccctcaa acagagacaa gaggactatt tggagctata gcagg 45

<210> 76
<211> 3435
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 676

<400> 76
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta ataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatggag aaaatagtgc ttcttcttgc aatagtcagt cttgttaaaa gtgatcagat 1320

ttgcattggt taccatgcaa acaattcaac agagcaggtt gacacaatca tggaaaagaa 1380

cgttactgtt acacatgccc aagacatact ggaaaagaca cacaacggga agctctgcga 1440

tctagatgga gtgaagcctc taattttaag agattgtagt gtagctggat ggctcctcgg 1500

gaacccaatg tgtgacgaat tcatcaatgt accggaatgg tcttacatag tggagaaggc 1560

caatccaacc aatgacctct gttacccagg gagtttcaac gactatgaag aactgaaaca 1620

cctattgagc agaataaacc attttgagaa aattcaaatc atccccaaaa gttcttggtc 1680

cgatcatgaa gcctcatcag gagttagctc agcatgtcca tacctgggaa gtccctcctt 1740

ttttagaaat gtggtatggc ttatcaaaaa gaacagtaca tacccaacaa taaagaaaag 1800

ctacaataat accaaccaag aggatctttt ggtactgtgg ggaattcacc atcctaatga 1860

tgcggcagag cagacaaggc tatatcaaaa cccaaccacc tatatttcca ttgggacatc 1920

aacactaaac cagagattgg taccaaaaat agctactaga tccaaagtaa acgggcaaag 1980

tggaaggatg gagttcttct ggacaatttt aaaacctaat gatgcaatca acttcgagag 2040

taatggaaat ttcattgctc cagaatatgc atacaaaatt gtcaagaaag gggactcagc 2100

aattatgaaa agtgaattgg aatatggtaa ctgcaacacc aagtgtcaaa ctccaatggg 2160

ggcgataaac tctagtatgc cattccacaa catacaccct ctcaccatcg gggaatgccc 2220

caaatatgtg aaatcaaaca gattagtcct tgcaacaggg ctcagaaata gccctcaaac 2280

agagacaaga ggactatttg gagctatagc aggttttata gagggaggat ggcagggaat 2340

ggtagatggt tggtatgggt accaccatag caatgagcag gggagtgggt acgctgcaga 2400

caaagaatcc actcaaaagg caatagatgg agtcaccaat aaggtcaact caatcattga 2460

caaaatgaac actcagtttg aggccgttgg aagggaattt aataacttag aaaggagaat 2520

agagaattta aacaagaaga tggaagacgg gtttctagat gtctggactt ataatgccga 2580

acttctggtt ctcatggaaa atgagagaac tctagacttt catgactcaa atgttaagaa 2640

cctctacgac aaggtccgac tacagcttag ggataatgca aaggagctgg gtaacggttg 2700

tttcgagttc tatcacaaat gtgataatga atgtatggaa agtataagaa acggaacgta 2760

caactatccg cagtattcag aagaagcaag attaaaaaga gaggaaataa gtggggtaaa 2820

attggaatca ataggaactt accaaatact gtcaatttat tcaacagtgg cgagttccct 2880

agcactggca atcatgatgg ctggtctatc tttatggatg tgctccaatg gatcgttaca 2940

atgcagaatt tgcatttaaa ggcctatttt ctttagtttg aatttactgt tattcggtgt 3000

gcatttctat gtttggtgag cggttttctg tgctcagagt gtgtttattt tatgtaattt 3060

aatttctttg tgagctcctg tttagcaggt cgtcccttca gcaaggacac aaaaagattt 3120

taattttatt aaaaaaaaaa aaaaaaaaga ccgggaattc gatatcaagc ttatcgacct 3180

gcagatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt 3240

gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa 3300

tgcatgacgt tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa 3360

tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca 3420

tctatgttac tagat 3435

<210> 77
<211> 564
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin cúm A/Indonesia/5/2005 (H5N1) TETR thể đột biến vị trí tách.

<400> 77

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
1 5 10 15

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20 25 30

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
35 40 45

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
50 55 60

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
65 70 75 80

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
85 90 95

```
Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn
            100                 105                 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
        115                 120                 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
    130                 135                 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
145                 150                 155                 160

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
                165                 170                 175

Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
            180                 185                 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
        195                 200                 205

Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
    210                 215                 220

Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
225                 230                 235                 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
                245                 250                 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
            260                 265                 270

Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
        275                 280                 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
    290                 295                 300
```

```
Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
305                 310                 315                 320

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
                325                 330                 335

Pro Gln Thr Glu Thr Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
            340                 345                 350

Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His
        355                 360                 365

Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln
    370                 375                 380

Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp Lys
385                 390                 395                 400

Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu
                405                 410                 415

Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp
            420                 425                 430

Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg
        435                 440                 445

Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val
    450                 455                 460

Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe
465                 470                 475                 480

Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Ile Arg Asn
                485                 490                 495

Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys Arg
            500                 505                 510

Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln Ile
        515                 520                 525
```

Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met
530 535 540

Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys
545 550 555 560

Arg Ile Cys Ile

<210> 78
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi H5I505_TETQ.r

<400> 78
tccaaatagt ccttgtgtct ctgtttgagg gctatttctg agccctgt 48

<210> 79
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi H5I505_TETQ.c

<400> 79
aaatagccct caaacagaga cacaaggact atttggagct atagcagg 48

<210> 80
<211> 3435
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 766

<400> 80
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatggag aaaatagtgc ttcttcttgc aatagtcagt cttgttaaaa gtgatcagat 1320

ttgcattggt taccatgcaa acaattcaac agagcaggtt gacacaatca tggaaaagaa 1380

cgttactgtt acacatgccc aagacatact ggaaaagaca cacaacggga agctctgcga 1440

tctagatgga gtgaagcctc taattttaag agattgtagt gtagctggat ggctcctcgg 1500

gaacccaatg tgtgacgaat tcatcaatgt accggaatgg tcttacatag tggagaaggc 1560

caatccaacc aatgacctct gttacccagg gagtttcaac gactatgaag aactgaaaca 1620

cctattgagc agaataaacc attttgagaa aattcaaatc atccccaaaa gttcttggtc 1680

cgatcatgaa gcctcatcag gagttagctc agcatgtcca tacctgggaa gtccctcctt 1740

ttttagaaat gtggtatggc ttatcaaaaa gaacagtaca tacccaacaa taaagaaaag 1800

ctacaataat accaaccaag aggatctttt ggtactgtgg ggaattcacc atcctaatga 1860

tgcggcagag cagacaaggc tatatcaaaa cccaaccacc tatatttcca ttgggacatc 1920

aacactaaac cagagattgg taccaaaaat agctactaga tccaaagtaa acgggcaaag 1980

tggaaggatg gagttcttct ggacaatttt aaaacctaat gatgcaatca acttcgagag 2040

taatggaaat ttcattgctc cagaatatgc atacaaaatt gtcaagaaag gggactcagc 2100

aattatgaaa agtgaattgg aatatggtaa ctgcaacacc aagtgtcaaa ctccaatggg 2160

ggcgataaac tctagtatgc cattccacaa catacaccct ctcaccatcg gggaatgccc 2220

caaatatgtg aaatcaaaca gattagtcct tgcaacaggg ctcagaaata gccctcaaac 2280

agagacacaa ggactatttg gagctatagc aggttttata gagggaggat ggcagggaat 2340

ggtagatggt tggtatgggt accaccatag caatgagcag gggagtgggt acgctgcaga 2400

caaagaatcc actcaaaagg caatagatgg agtcaccaat aaggtcaact caatcattga 2460

caaaatgaac actcagtttg aggccgttgg aagggaattt aataacttag aaaggagaat 2520

agagaattta aacaagaaga tggaagacgg gtttctagat gtctggactt ataatgccga 2580

acttctggtt ctcatggaaa atgagagaac tctagacttt catgactcaa atgttaagaa 2640

cctctacgac aaggtccgac tacagcttag ggataatgca aaggagctgg gtaacggttg 2700

tttcgagttc tatcacaaat gtgataatga atgtatggaa agtataagaa acggaacgta 2760

caactatccg cagtattcag aagaagcaag attaaaaaga gaggaaataa gtggggtaaa 2820

attggaatca ataggaactt accaaatact gtcaatttat tcaacagtgg cgagttccct 2880

agcactggca atcatgatgg ctggtctatc tttatggatg tgctccaatg gatcgttaca 2940

atgcagaatt tgcatttaaa ggcctatttt ctttagtttg aatttactgt tattcggtgt 3000

gcatttctat gtttggtgag cggttttctg tgctcagagt gtgtttattt tatgtaattt 3060

aatttctttg tgagctcctg tttagcaggt cgtcccttca gcaaggacac aaaaagattt 3120

taattttatt aaaaaaaaaa aaaaaaaaga ccgggaattc gatatcaagc ttatcgacct 3180

gcagatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt 3240

gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa 3300

tgcatgacgt tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa 3360

tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca 3420

tctatgttac tagat 3435

<210> 81
<211> 564
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin cúm A/Indonesia/5/2005 (H5N1) TETQ thể đột biến vị trí tách

<400> 81

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
1 5 10 15

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20 25 30

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
35 40 45

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
50 55 60

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
65 70 75 80

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
85 90 95

Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn
100 105 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
115 120 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
130 135 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
145 150 155 160

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile

```
                 165                 170                 175

Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
            180                 185                 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
        195                 200                 205

Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
    210                 215                 220

Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
225                 230                 235                 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
                245                 250                 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
            260                 265                 270

Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
        275                 280                 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
    290                 295                 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
305                 310                 315                 320

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
                325                 330                 335

Pro Gln Thr Glu Thr Gln Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile
            340                 345                 350

Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His
        355                 360                 365

Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln
    370                 375                 380
```

Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp Lys
385 390 395 400

Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu
405 410 415

Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp
420 425 430

Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg
435 440 445

Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val
450 455 460

Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe
465 470 475 480

Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Ile Arg Asn
485 490 495

Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys Arg
500 505 510

Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln Ile
515 520 525

Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met
530 535 540

Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys
545 550 555 560

Arg Ile Cys Ile

<210> 82
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi H5I505(PrL-).r

<400> 82
ctgccatcct ccgccagggc tatttctgag ccctgttgca aggactaatc 50

<210> 83
<211> 52
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mồi H5I505(PrL-).c

<400> 83
gaaatagccc tggcggagga tggcagggaa tggtagatgg ttggtatggg ta 52

<210> 84
<211> 3390
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 928

<400> 84
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatggag aaaatagtgc ttcttcttgc aatagtcagt cttgttaaaa gtgatcagat 1320

ttgcattggt taccatgcaa acaattcaac agagcaggtt gacacaatca tggaaaagaa 1380

cgttactgtt acacatgccc aagacatact ggaaaagaca cacaacggga agctctgcga 1440

tctagatgga gtgaagcctc taattttaag agattgtagt gtagctggat ggctcctcgg 1500

gaacccaatg tgtgacgaat tcatcaatgt accggaatgg tcttacatag tggagaaggc 1560

caatccaacc aatgacctct gttacccagg gagtttcaac gactatgaag aactgaaaca 1620

cctattgagc agaataaacc attttgagaa aattcaaatc atccccaaaa gttcttggtc 1680

cgatcatgaa gcctcatcag gagttagctc agcatgtcca tacctgggaa gtccctcctt 1740

ttttagaaat gtggtatggc ttatcaaaaa gaacagtaca tacccaacaa taaagaaaag 1800

ctacaataat accaaccaag aggatctttt ggtactgtgg ggaattcacc atcctaatga 1860

tgcggcagag cagacaaggc tatatcaaaa cccaaccacc tatatttcca ttgggacatc 1920

aacactaaac cagagattgg taccaaaaat agctactaga tccaaagtaa acgggcaaag 1980

tggaaggatg gagttcttct ggacaatttt aaaacctaat gatgcaatca acttcgagag 2040

taatggaaat ttcattgctc cagaatatgc atacaaaatt gtcaagaaag gggactcagc 2100

aattatgaaa agtgaattgg aatatggtaa ctgcaacacc aagtgtcaaa ctccaatggg 2160

ggcgataaac tctagtatgc cattccacaa catacaccct ctcaccatcg gggaatgccc 2220

caaatatgtg aaatcaaaca gattagtcct tgcaacaggg ctcagaaata gccctggcgg 2280

aggatggcag ggaatggtag atggttggta tgggtaccac catagcaatg agcaggggag 2340

tgggtacgct gcagacaaag aatccactca aaaggcaata gatggagtca ccaataaggt 2400

caactcaatc attgacaaaa tgaacactca gtttgaggcc gttggaaggg aatttaataa 2460

cttagaaagg agaatagaga atttaaacaa gaagatggaa gacgggtttc tagatgtctg 2520

gacttataat gccgaacttc tggttctcat ggaaaatgag agaactctag actttcatga 2580

ctcaaatgtt aagaacctct acgacaaggt ccgactacag cttagggata atgcaaagga 2640

gctgggtaac ggttgtttcg agttctatca caaatgtgat aatgaatgta tggaaagtat 2700

aagaaacgga acgtacaact atccgcagta ttcagaagaa gcaagattaa aaagagagga 2760

aataagtggg gtaaaattgg aatcaatagg aacttaccaa atactgtcaa tttattcaac 2820

agtggcgagt tccctagcac tggcaatcat gatggctggt ctatctttat ggatgtgctc 2880

caatggatcg ttacaatgca gaatttgcat ttaaaggcct attttcttta gtttgaattt 2940

actgttattc ggtgtgcatt tctatgtttg gtgagcggtt ttctgtgctc agagtgtgtt 3000

tattttatgt aatttaattt ctttgtgagc tcctgtttag caggtcgtcc cttcagcaag 3060

gacacaaaaa gattttaatt ttattaaaaa aaaaaaaaaa aaagaccggg aattcgatat 3120

caagcttatc gacctgcaga tcgttcaaac atttggcaat aaagtttctt aagattgaat 3180

cctgttgccg gtcttgcgat gattatcata taatttctgt tgaattacgt taagcatgta 3240

ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat tagagtcccg 3300

caattataca tttaatacgc gatagaaaac aaaatatagc gcgcaaacta ggataaatta 3360

tcgcgcgcgg tgtcatctat gttactagat 3390

<210> 85
<211> 549
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin cúm A/Indonesia/5/2005 (H5N1) với vòng phân giải protein loại bớt.

<400> 85

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
1 5 10 15

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20 25 30

```
Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
        35                  40                  45

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
    50                  55                  60

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
65                  70                  75                  80

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
                85                  90                  95

Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn
            100                 105                 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
        115                 120                 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
    130                 135                 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
145                 150                 155                 160

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
                165                 170                 175

Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
            180                 185                 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
        195                 200                 205

Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
    210                 215                 220

Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
225                 230                 235                 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
                245                 250                 255
```

```
Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
            260                 265                 270

Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
        275                 280                 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
    290                 295                 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
305                 310                 315                 320

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
                325                 330                 335

Pro Gly Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His
            340                 345                 350

His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr
        355                 360                 365

Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp
    370                 375                 380

Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu
385                 390                 395                 400

Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu
                405                 410                 415

Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu
            420                 425                 430

Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys
        435                 440                 445

Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys
    450                 455                 460
```

Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Ile Arg
465 470 475 480

Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys
485 490 495

Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln
500 505 510

Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile
515 520 525

Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln
530 535 540

Cys Arg Ile Cys Ile
545

<210> 86
<211> 40
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1029 IF-S2+S4-B Bris.c

<400> 86
tctcagatct tcgccgatcg aatctgcact ggaataacat 40

<210> 87
<211> 43
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1029 IF-S1a4-B Bris.r

<400> 87
actaaagaaa ataggccttt atagacagat ggagcaagaa aca 43

<210> 88
<211> 1758
<212> ADN
<213> Trình tự nhân tạo

<220>

<223> gen HA B Brisbane tổng hợp

<400> 88

atgaaggcaa taattgtact actcatggta gtaacatcca atgcagatcg aatctgcact 60

ggaataacat cgtcaaactc accacatgtc gtcaaaactg ctactcaagg ggaggtcaat 120

gtgactggtg taataccact gacaacaaca cccaccaaat ctcattttgc aaatctcaaa 180

ggaacagaaa ccaggggggaa actatgccca aaatgcctca actgcacaga tctggacgta 240

gccttgggca gaccaaaatg cacggggaaa ataccctcgg caagagtttc aatactccat 300

gaagtcagac ctgttacatc tgggtgcttt cctataatgc acgacagaac aaaaattaga 360

cagctgccta accttctccg aggatacgaa catatcaggt tatcaaccca taacgttatc 420

aatgcagaaa atgcaccagg aggaccctac aaaattggaa cctcagggtc ttgccctaac 480

attaccaatg gaaacggatt tttcgcaaca atggcttggg ccgtcccaaa aaacgacaaa 540

aacaaaacag caacaaatcc attaacaata gaagtaccat acatttgtac agaaggagaa 600

gaccaaatta ccgtttgggg gttccactct gacaacgaga cccaaatggc aaagctctat 660

ggggactcaa agccccagaa gttcacctca tctgccaacg gagtgaccac acattacgtt 720

tcacagattg gtggcttccc aaatcaaaca gaagacggag gactaccaca aagtggtaga 780

attgttgttg attacatggt gcaaaaatct gggaaaacag gaacaattac ctatcaaagg 840

ggtattttat tgcctcaaaa ggtgtggtgc gcaagtggca ggagcaaggt aataaaagga 900

tccttgcctt taattggaga agcagattgc ctccacgaaa aatacggtgg attaaacaaa 960

agcaagcctt actacacagg ggaacatgca aaggccatag gaaattgccc aatatgggtg 1020

aaaacaccct tgaagctggc caatggaacc aaatatagac ctcctgcaaa actattaaag 1080

gaaaggggtt tcttcggagc tattgctggt ttcttagaag gaggatggga aggaatgatt 1140

gcaggttggc acggatacac atcccatggg gcacatggag tagcggtggc agcagacctt 1200

aagagcactc aagaggccat aaacaagata acaaaaaatc tcaactcttt gagtgagctg 1260

gaagtaaaga atcttcaaag actaagcggt gccatggatg aactccacaa cgaaatacta 1320

gaactagatg agaaagtgga tgatctcaga gctgatacaa taagctcaca aatagaactc 1380

gcagtcctgc tttccaatga aggaataata aacagtgaag atgaacatct cttggcgctt 1440

gaaagaaagc tgaagaaaat gctgggcccc tctgctgtag agatagggaa tggatgcttt 1500

gaaaccaaac acaagtgcaa ccagacctgt ctcgacagaa tagctgctgg taccttttgat 1560

gcaggagaat tttctctccc cacctttgat tcactgaata ttactgctgc atctttaaat 1620

gacgatggat tggataatca tactatactg ctttactact caactgctgc ctccagtttg 1680

gctgtaacac tgatgatagc tatctttgtt gtttatatgg tctccagaga caatgtttct 1740

tgctccatct gtctataa 1758

<210> 89
<211> 3525
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 1029

<400> 89
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta ataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccatggcg aaaaacgttg cgattttcgg cttattgttt tctcttcttg tgttggttcc 1320

ttctcagatc ttcgccgatc gaatctgcac tggaataaca tcgtcaaact caccacatgt 1380

cgtcaaaact gctactcaag gggaggtcaa tgtgactggt gtaataccac tgacaacaac 1440

acccaccaaa tctcattttg caaatctcaa aggaacagaa accaggggga aactatgccc 1500

aaaatgcctc aactgcacag atctggacgt agccttgggc agaccaaaat gcacggggaa 1560

aataccctcg gcaagagttt caatactcca tgaagtcaga cctgttacat ctgggtgctt 1620

tcctataatg cacgacagaa caaaaattag acagctgcct aaccttctcc gaggatacga 1680

acatatcagg ttatcaaccc ataacgttat caatgcagaa aatgcaccag gaggacccta 1740

caaaattgga acctcagggt cttgccctaa cattaccaat ggaaacggat ttttcgcaac 1800

aatggcttgg gccgtcccaa aaaacgacaa aaacaaaaca gcaacaaatc cattaacaat 1860

agaagtacca tacatttgta cagaaggaga agaccaaatt accgtttggg ggttccactc 1920

tgacaacgag acccaaatgg caaagctcta tggggactca aagccccaga agttcacctc 1980

atctgccaac ggagtgacca cacattacgt ttcacagatt ggtggcttcc caaatcaaac 2040

agaagacgga ggactaccac aaagtggtag aattgttgtt gattacatgg tgcaaaaatc 2100

tgggaaaaca ggaacaatta cctatcaaag gggtatttta ttgcctcaaa aggtgtggtg 2160

cgcaagtggc aggagcaagg taataaaagg atccttgcct ttaattggag aagcagattg 2220

cctccacgaa aaatacggtg gattaaacaa aagcaagcct tactacacag gggaacatgc 2280

aaaggccata ggaaattgcc caatatgggt gaaaacaccc ttgaagctgg ccaatggaac 2340

caaatataga cctcctgcaa aactattaaa ggaaaggggt ttcttcggag ctattgctgg 2400

tttcttagaa ggaggatggg aaggaatgat tgcaggttgg cacggataca catcccatgg 2460

ggcacatgga gtagcggtgg cagcagacct taagagcact caagaggcca taaacaagat 2520

aacaaaaaat ctcaactctt tgagtgagct ggaagtaaag aatcttcaaa gactaagcgg 2580

tgccatggat gaactccaca acgaaatact agaactagat gagaaagtgg atgatctcag 2640

agctgataca ataagctcac aaatagaact cgcagtcctg ctttccaatg aaggaataat 2700

aaacagtgaa gatgaacatc tcttggcgct tgaaagaaag ctgaagaaaa tgctgggccc 2760

ctctgctgta gagataggga atggatgctt tgaaaccaaa cacaagtgca accagacctg 2820

tctcgacaga atagctgctg gtacctttga tgcaggagaa ttttctctcc ccacctttga 2880

ttcactgaat attactgctg catctttaaa tgacgatgga ttggataatc atactatact 2940

gctttactac tcaactgctg cctccagttt ggctgtaaca ctgatgatag ctatctttgt 3000

tgtttatatg gtctccagag acaatgtttc ttgctccatc tgtctataaa ggcctatttt 3060

ctttagtttg aatttactgt tattcggtgt gcatttctat gtttggtgag cggttttctg 3120

tgctcagagt gtgtttattt tatgtaattt aatttctttg tgagctcctg tttagcaggt 3180

cgtcccttca gcaaggacac aaaaagattt taattttatt aaaaaaaaaa aaaaaaaaga 3240

ccgggaattc gatatcaagc ttatcgacct gcagatcgtt caaacatttg gcaataaagt 3300

ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 3360

tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggttttt 3420

atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 3480

aactaggata aattatcgcg cgcggtgtca tctatgttac tagat 3525

<210> 90
<211> 594
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA từ cúm B/Brisbane/60/2008

<400> 90

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
50 55 60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
85 90 95

Thr Gly Lys Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
115 120 125

Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
130 135 140

Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
145 150 155 160

Ile Gly Thr Ser Gly Ser Cys Pro Asn Ile Thr Asn Gly Asn Gly Phe
165 170 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Lys Asn Lys Thr
180 185 190

Ala Thr Asn Pro Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly
195 200 205

Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu Thr Gln
210 215 220

Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser
225 230 235 240

Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro
245 250 255

Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val
260 265 270

```
Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln
        275                 280                 285

Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser
    290                 295                 300

Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu
305                 310                 315                 320

His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly
                325                 330                 335

Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro
            340                 345                 350

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys Leu Leu
        355                 360                 365

Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu Gly Gly
    370                 375                 380

Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala
385                 390                 395                 400

His Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile
                405                 410                 415

Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys
            420                 425                 430

Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile
        435                 440                 445

Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser
    450                 455                 460

Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn
465                 470                 475                 480

Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met
                485                 490                 495
```

Leu Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys
500 505 510

His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe
515 520 525

Asp Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr
530 535 540

Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu
545 550 555 560

Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Ile Ala
565 570 575

Ile Phe Val Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile
580 585 590

Cys Leu

<210> 91
<211> 1734
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Brisbane (PrL-).

<400> 91
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg ccgatcgaat ctgcactgga ataacatcgt caaactcacc acatgtcgtc 120

aaaactgcta ctcaagggga ggtcaatgtg actggtgtaa taccactgac aacaacaccc 180

accaaatctc attttgcaaa tctcaaagga acagaaacca gggggaaact atgcccaaaa 240

tgcctcaact gcacagatct ggacgtagcc ttgggcagac caaaatgcac ggggaaaata 300

ccctcggcaa gagtttcaat actccatgaa gtcagacctg ttacatctgg gtgctttcct 360

ataatgcacg acagaacaaa aattagacag ctgcctaacc ttctccgagg atacgaacat 420

atcaggttat caacccataa cgttatcaat gcagaaaatg caccaggagg accctacaaa 480

attggaacct cagggtcttg ccctaacatt accaatggaa acggattttt cgcaacaatg 540

gcttgggccg tcccaaaaaa cgacaaaaac aaaacagcaa caaatccatt aacaatagaa 600

gtaccataca tttgtacaga aggagaagac caaattaccg tttgggggtt ccactctgac 660

aacgagaccc aaatggcaaa gctctatggg gactcaaagc cccagaagtt cacctcatct 720

gccaacggag tgaccacaca ttacgtttca cagattggtg gcttcccaaa tcaaacagaa 780

gacggaggac taccacaaag tggtagaatt gttgttgatt acatggtgca aaaatctggg 840

aaaacaggaa caattaccta tcaaaggggt attttattgc ctcaaaaggt gtggtgcgca 900

agtggcagga gcaaggtaat aaaaggatcc ttgcctttaa ttggagaagc agattgcctc 960

cacgaaaaat acggtggatt aaacaaaagc aagccttact acacagggga acatgcaaag 1020

gccataggaa attgcccaat atgggtgaaa acacccttga agctggccaa tggaaccaaa 1080

tatagacctc ctggtggagg atgggaagga atgattgcag gttggcacgg atacacatcc 1140

catggggcac atggagtagc ggtggcagca gaccttaaga gcactcaaga ggccataaac 1200

aagataacaa aaaatctcaa ctctttgagt gagctggaag taaagaatct tcaaagacta 1260

agcggtgcca tggatgaact ccacaacgaa atactagaac tagatgagaa agtggatgat 1320

ctcagagctg atacaataag ctcacaaata gaactcgcag tcctgctttc caatgaagga 1380

ataataaaca gtgaagatga acatctcttg gcgcttgaaa gaaagctgaa gaaaatgctg 1440

ggcccctctg ctgtagagat agggaatgga tgctttgaaa ccaaacacaa gtgcaaccag 1500

acctgtctcg acagaatagc tgctggtacc tttgatgcag gagaattttc tctccccacc 1560

tttgattcac tgaatattac tgctgcatct ttaaatgacg atggattgga taatcatact 1620

atactgcttt actactcaac tgctgcctcc agtttggctg taacactgat gatagctatc 1680

tttgttgttt atatggtctc cagagacaat gtttcttgct ccatctgtct ataa 1734

<210> 92
<211> 577
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Brisbane (PrL-)

<400> 92

```
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                  10                  15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
            20                  25                  30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
        35                  40                  45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
    50                  55                  60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
65                  70                  75                  80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
                85                  90                  95

Thr Gly Lys Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
            100                 105                 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
        115                 120                 125

Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
    130                 135                 140

Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
145                 150                 155                 160

Ile Gly Thr Ser Gly Ser Cys Pro Asn Ile Thr Asn Gly Asn Gly Phe
                165                 170                 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Lys Asn Lys Thr
            180                 185                 190

Ala Thr Asn Pro Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly
        195                 200                 205

Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu Thr Gln
    210                 215                 220
```

Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser
225 230 235 240

Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro
245 250 255

Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val
260 265 270

Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln
275 280 285

Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser
290 295 300

Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu
305 310 315 320

His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly
325 330 335

Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro
340 345 350

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp
355 360 365

Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His
370 375 380

Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn
385 390 395 400

Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn
405 410 415

Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu
420 425 430

Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser
435 440 445

Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser
450 455 460

Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu
465 470 475 480

Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys His
485 490 495

Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asp
500 505 510

Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala
515 520 525

Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu Tyr
530 535 540

Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Ile Ala Ile
545 550 555 560

Phe Val Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile Cys
565 570 575

Leu

<210> 93
<211> 160
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn tăng cường sự biểu hiện CPMVX

<400> 93
tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc 60

ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc ttgcgtgagc 120

gatcttcaac gttgtcagat cgtgcttcgg caccagtaca 160

<210> 94
<211> 29
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin thống nhất của màng chuyển qua màng

<220>
<221> misc_dấu hiệu
<222> (3)..(3)
<223> Xaa có thể là axit amin xuất hiện tự nhiên bất kỳ

<220>
<221> misc_dấu hiệu
<222> (15)..(15)
<223> Xaa có thể là axit amin xuất hiện tự nhiên bất kỳ

<220>
<221> misc_dấu hiệu
<222> (17)..(18)
<223> Xaa có thể là axit amin xuất hiện tự nhiên bất kỳ

<220>
<221> misc_dấu hiệu
<222> (23)..(23)
<223> Xaa có thể là axit amin xuất hiện tự nhiên bất kỳ

<220>
<221> misc_dấu hiệu
<222> (25)..(25)
<223> Xaa có thể là axit amin xuất hiện tự nhiên bất kỳ

<400> 94

Ile Leu Xaa Ile Tyr Tyr Ser Thr Val Ala Ile Ser Ser Leu Xaa Leu
1 5 10 15

Xaa Xaa Met Leu Ala Gly Xaa Ser Xaa Trp Met Cys Ser
20 25

<210> 95
<211> 1734
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Brisbane (PrL-)+H1 California TMCT.

<400> 95
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg ccgatcgaat ctgcactgga ataacatcgt caaactcacc acatgtcgtc 120

aaaactgcta ctcaagggga ggtcaatgtg actggtgtaa taccactgac aacaacaccc 180

accaaatctc attttgcaaa tctcaaagga acagaaacca gggggaaact atgcccaaaa 240

tgcctcaact gcacagatct ggacgtagcc ttgggcagac caaaatgcac ggggaaaata 300

ccctcggcaa gagtttcaat actccatgaa gtcagacctg ttacatctgg gtgctttcct 360

ataatgcacg acagaacaaa aattagacag ctgcctaacc ttctccgagg atacgaacat 420

atcaggttat caacccataa cgttatcaat gcagaaaatg caccaggagg accctacaaa 480

attggaacct cagggtcttg ccctaacatt accaatggaa acggattttt cgcaacaatg 540

gcttgggccg tcccaaaaaa cgacaaaaac aaaacagcaa caaatccatt aacaatagaa 600

gtaccataca tttgtacaga aggagaagac caaattaccg tttgggggtt ccactctgac 660

aacgagaccc aaatggcaaa gctctatggg gactcaaagc cccagaagtt cacctcatct 720

gccaacggag tgaccacaca ttacgtttca cagattggtg gcttcccaaa tcaaacagaa 780

gacggaggac taccacaaag tggtagaatt gttgttgatt acatggtgca aaaatctggg 840

aaaacaggaa caattaccta tcaaaggggt attttattgc ctcaaaaggt gtggtgcgca 900

agtggcagga gcaaggtaat aaaaggatcc ttgcctttaa ttggagaagc agattgcctc 960

cacgaaaaat acggtggatt aaacaaaagc aagccttact acacagggga acatgcaaag 1020

gccataggaa attgcccaat atgggtgaaa acacccttga agctggccaa tggaaccaaa 1080

tatagacctc ctggtggagg atgggaagga atgattgcag gttggcacgg atacacatcc 1140

catggggcac atggagtagc ggtggcagca gaccttaaga gcactcaaga ggccataaac 1200

aagataacaa aaaatctcaa ctctttgagt gagctggaag taaagaatct tcaaagacta 1260

agcggtgcca tggatgaact ccacaacgaa atactagaac tagatgagaa agtggatgat 1320

ctcagagctg atacaataag ctcacaaata gaactcgcag tcctgctttc caatgaagga 1380

ataataaaca gtgaagatga acatctcttg gcgcttgaaa gaaagctgaa gaaaatgctg 1440

ggcccctctg ctgtagagat agggaatgga tgctttgaaa ccaaacacaa gtgcaaccag 1500

acctgtctcg acagaatagc tgctggtacc tttgatgcag gagaattttc tctccccacc 1560

tttgattcac tgaatattac tgctgcatct ttaaatgacg atggattgga taattaccag 1620

attttggcga tctattcaac tgtcgccagt tcattggtac tggtagtctc cctgggggca 1680

atcagtttct ggatgtgctc taatgggtct ctacagtgta gaatatgtat ttaa . 1734

<210> 96
<211> 577
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Brisbane (PrL-)+H1 California TMCT

<400> 96

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
50 55 60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
85 90 95

Thr Gly Lys Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
115 120 125

Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
130 135 140

```
Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
145                 150                 155                 160

Ile Gly Thr Ser Gly Ser Cys Pro Asn Ile Thr Asn Gly Asn Gly Phe
                165                 170                 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Lys Asn Lys Thr
            180                 185                 190

Ala Thr Asn Pro Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly
        195                 200                 205

Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu Thr Gln
    210                 215                 220

Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser
225                 230                 235                 240

Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro
                245                 250                 255

Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val
            260                 265                 270

Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln
        275                 280                 285

Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser
    290                 295                 300

Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu
305                 310                 315                 320

His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly
                325                 330                 335

Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro
            340                 345                 350

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp
        355                 360                 365
```

Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His
370 375 380

Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn
385 390 395 400

Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn
405 410 415

Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu
420 425 430

Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser
435 440 445

Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser
450 455 460

Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu
465 470 475 480

Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys His
485 490 495

Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asp
500 505 510

Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala
515 520 525

Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn Tyr Gln Ile Leu Ala Ile
530 535 540

Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Val Val Ser Leu Gly Ala
545 550 555 560

Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys
565 570 575

Ile

<210> 97
<211> 1731
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-)

<400> 97
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg ccgatcgaat ctgcactggg ataacatctt caaactcacc tcatgtggtc 120

aaaacagcta ctcaagggga ggtcaatgtg actggtgtga taccactaac aacaacacca 180

acaaaatctt attttgcaaa tctcaaagga acaaagacca gagggaaact atgcccagac 240

tgtctcaact gtacagatct ggatgtggcc ctgggcaggc caatgtgtgt gggaactaca 300

ccttctgcga aagcttcaat acttcacgaa gtcagacctg ttacatccgg gtgcttccct 360

ataatgcacg acagaacaaa aatcaggcaa ctagccaatc ttctcagagg atatgaaaat 420

atcaggttat caacccaaaa cgttatcgat gcagaaaagg caccaggagg accctacaga 480

cttggaacct caggatcttg ccctaacgct accagtaaaa gcggattttt cgcaacaatg 540

gcttgggctg tcccaaagga caacaacaaa aatgcaacga acccattaac agtagaagta 600

ccatacattt gtgcagaagg ggaagaccaa attactgttt gggggttcca ttcagataac 660

aaaacccaaa tgaagaacct ctatggagac tcaaatcctc aaaagttcac ctcatctgct 720

aatggagtaa ccacacatta tgtttctcag attggcggct tcccagatca aacagaagac 780

ggaggactac cacaaagcgg cagaattgtc gttgattaca tgatgcaaaa acctgggaaa 840

acaggaacaa ttgtctatca aagaggtgtt ttgttgcctc aaaaggtgtg gtgcgcgagt 900

ggcaggagca aagtaataaa agggtccttg cctttaattg gtgaagcaga ttgccttcat 960

gaaaaatacg gtggattaaa caaaagcaag ccttactaca caggagaaca tgcaaaagcc 1020

ataggaaatt gcccaatatg ggtgaaaaca cctttgaagc ttgccaatgg aaccaaatat 1080

agacctcctg gtggaggatg ggaaggaatg attgcaggtt ggcacggata cacatctcac 1140

ggagcacatg gagtggcagt tgctgcagac cttaagagca cacaagaagc tataaacaag 1200

ataacaaaaa atctcaactc tttgagtgag ctagaagtaa agaatcttca aaggctaagt 1260

ggtgccatgg atgaactcca caacgaaata ctcgagctgg atgagaaagt ggatgacctc 1320

agagctgaca ctataagttc acaaatagaa cttgcagtct tgctttccaa cgaaggaata 1380

ataaacagtg aagacgagca tctattggca cttgagagaa aactaaagaa aatgctgggt 1440

ccctctgctg tagacatagg aaatggatgc ttcgaaacca aacacaaatg caaccagacc 1500

tgcttagaca ggatagctgc tggcaccttt aatgcaggag agttttctct ccccactttt 1560

gattcattga acattactgc tgcatcttta aatgatgatg gattggataa ccatactata 1620

ctgctctatt actcaactgc tgcttctagt ttggctgtaa cattgatgct agctattttt 1680

attgtttata tggtctccag agacaacgtt tcatgctcca tctgtctata a 1731

<210> 98
<211> 576
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Massachussetts (PrL-)

<400> 98

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser Tyr
50 55 60

Phe Ala Asn Leu Lys Gly Thr Lys Thr Arg Gly Lys Leu Cys Pro Asp
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Met Cys
85 90 95

Val Gly Thr Thr Pro Ser Ala Lys Ala Ser Ile Leu His Glu Val Arg
100 105 110

```
Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
        115                 120                 125

Arg Gln Leu Ala Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser
    130                 135                 140

Thr Gln Asn Val Ile Asp Ala Glu Lys Ala Pro Gly Gly Pro Tyr Arg
145                 150                 155                 160

Leu Gly Thr Ser Gly Ser Cys Pro Asn Ala Thr Ser Lys Ser Gly Phe
                165                 170                 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asp Asn Asn Lys Asn Ala
            180                 185                 190

Thr Asn Pro Leu Thr Val Glu Val Pro Tyr Ile Cys Ala Glu Gly Glu
        195                 200                 205

Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Lys Thr Gln Met
    210                 215                 220

Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala
225                 230                 235                 240

Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro Asp
                245                 250                 255

Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp
            260                 265                 270

Tyr Met Met Gln Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg
        275                 280                 285

Gly Val Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys
    290                 295                 300

Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu His
305                 310                 315                 320

Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly Glu
```

325 330 335

His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro Leu
340 345 350

Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp Glu
355 360 365

Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His Gly
370 375 380

Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn Lys
385 390 395 400

Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn Leu
405 410 415

Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu Glu
420 425 430

Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser Gln
435 440 445

Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser Glu
450 455 460

Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu Gly
465 470 475 480

Pro Ser Ala Val Asp Ile Gly Asn Gly Cys Phe Glu Thr Lys His Lys
485 490 495

Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asn Ala
500 505 510

Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala Ala
515 520 525

Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu Tyr Tyr
530 535 540

Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Leu Ala Ile Phe
545 550 555 560

Ile Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile Cys Leu
565 570 575

<210> 99
<211> 1731
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT

<400> 99
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg ccgatcgaat ctgcactggg ataacatctt caaactcacc tcatgtggtc 120

aaaacagcta ctcaagggga ggtcaatgtg actggtgtga taccactaac aacaacacca 180

acaaaatctt attttgcaaa tctcaaagga acaaagacca gagggaaact atgcccagac 240

tgtctcaact gtacagatct ggatgtggcc ctgggcaggc caatgtgtgt gggaactaca 300

ccttctgcga aagcttcaat acttcacgaa gtcagacctg ttacatccgg gtgcttccct 360

ataatgcacg acagaacaaa aatcaggcaa ctagccaatc ttctcagagg atatgaaaat 420

atcaggttat caacccaaaa cgttatcgat gcagaaaagg caccaggagg accctacaga 480

cttggaacct caggatcttg ccctaacgct accagtaaaa gcggattttt cgcaacaatg 540

gcttgggctg tcccaaagga caacaacaaa aatgcaacga acccattaac agtagaagta 600

ccatacattt gtgcagaagg ggaagaccaa attactgttt gggggttcca ttcagataac 660

aaaacccaaa tgaagaacct ctatggagac tcaaatcctc aaaagttcac ctcatctgct 720

aatggagtaa ccacacatta tgtttctcag attggcggct tcccagatca aacagaagac 780

ggaggactac cacaaagcgg cagaattgtc gttgattaca tgatgcaaaa acctgggaaa 840

acaggaacaa ttgtctatca aagaggtgtt ttgttgcctc aaaaggtgtg gtgcgcgagt 900

ggcaggagca aagtaataaa agggtccttg cctttaattg gtgaagcaga ttgccttcat 960

gaaaaatacg gtggattaaa caaaagcaag ccttactaca caggagaaca tgcaaaagcc 1020

ataggaaatt gcccaatatg ggtgaaaaca cctttgaagc ttgccaatgg aaccaaatat 1080

agacctcctg gtggaggatg ggaaggaatg attgcaggtt ggcacggata cacatctcac 1140

ggagcacatg gagtggcagt tgctgcagac cttaagagca cacaagaagc tataaacaag 1200

ataacaaaaa atctcaactc tttgagtgag ctagaagtaa agaatcttca aaggctaagt 1260

ggtgccatgg atgaactcca caacgaaata ctcgagctgg atgagaaagt ggatgacctc 1320

agagctgaca ctataagttc acaaatagaa cttgcagtct tgctttccaa cgaaggaata 1380

ataaacagtg aagacgagca tctattggca cttgagagaa aactaaagaa aatgctgggt 1440

ccctctgctg tagacatagg aaatggatgc ttcgaaacca aacacaaatg caaccagacc 1500

tgcttagaca ggatagctgc tggcaccttt aatgcaggag agttttctct ccccactttt 1560

gattcattga acattactgc tgcatcttta aatgatgatg gattggataa ctaccagatt 1620

ttggcgatct attcaactgt cgccagttca ttggtactgg tagtctccct gggggcaatc 1680

agtttctgga tgtgctctaa tgggtctcta cagtgtagaa tatgtattta a 1731

<210> 100
<211> 576
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT

<400> 100

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser Tyr
50 55 60

Phe Ala Asn Leu Lys Gly Thr Lys Thr Arg Gly Lys Leu Cys Pro Asp
65 70 75 80

```
Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Met Cys
                85                  90                  95

Val Gly Thr Thr Pro Ser Ala Lys Ala Ser Ile Leu His Glu Val Arg
            100                 105                 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
        115                 120                 125

Arg Gln Leu Ala Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser
    130                 135                 140

Thr Gln Asn Val Ile Asp Ala Glu Lys Ala Pro Gly Gly Pro Tyr Arg
145                 150                 155                 160

Leu Gly Thr Ser Gly Ser Cys Pro Asn Ala Thr Ser Lys Ser Gly Phe
                165                 170                 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asp Asn Asn Lys Asn Ala
            180                 185                 190

Thr Asn Pro Leu Thr Val Glu Val Pro Tyr Ile Cys Ala Glu Gly Glu
        195                 200                 205

Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Lys Thr Gln Met
    210                 215                 220

Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala
225                 230                 235                 240

Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro Asp
                245                 250                 255

Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp
            260                 265                 270

Tyr Met Met Gln Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg
        275                 280                 285

Gly Val Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys
```

290 295 300

Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu His
305 310 315 320

Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly Glu
325 330 335

His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro Leu
340 345 350

Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp Glu
355 360 365

Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His Gly
370 375 380

Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn Lys
385 390 395 400

Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn Leu
405 410 415

Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu Glu
420 425 430

Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser Gln
435 440 445

Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser Glu
450 455 460

Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu Gly
465 470 475 480

Pro Ser Ala Val Asp Ile Gly Asn Gly Cys Phe Glu Thr Lys His Lys
485 490 495

Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asn Ala
500 505 510

Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala Ala
515 520 525

Ser Leu Asn Asp Asp Gly Leu Asp Asn Tyr Gln Ile Leu Ala Ile Tyr
530 535 540

Ser Thr Val Ala Ser Ser Leu Val Leu Val Val Ser Leu Gly Ala Ile
545 550 555 560

Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
565 570 575

<210> 101
<211> 1704
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của HA B Wisconsin (PrL-).

<400> 101
atgaaggcaa taattgtact actcatggta gtaacatcca atgcagatcg aatctgcact 60

gggataacat cttcaaactc acctcatgtg gtcaaaacag ctactcaagg ggaggtcaat 120

gtgactggcg tgataccact gacaacaaca ccaacaaaat cttattttgc aaatctcaaa 180

ggaacaagga ccagagggaa actatgcccg gactgtctca actgtacaga tctggatgtg 240

gccttgggca ggccaatgtg tgtggggacc acaccttctg ctaaagcttc aatactccac 300

gaggtcagac ctgttacatc cgggtgcttt cctataatgc acgacagaac aaaaatcagg 360

caactaccca atcttctcag aggatatgaa aatatcaggt tatcaaccca aaacgttatc 420

gatgcagaaa aagcaccagg aggaccctac agacttggaa cctcaggatc ttgccctaac 480

gctaccagta aaatcggatt ttttgcaaca atggcttggg ctgtcccaaa ggacaactac 540

aaaaatgcaa cgaacccact aacagtagaa gtaccataca tttgtacaga aggggaagac 600

caaattactg tttgggggtt ccattcagat aacaaaaccc aaatgaagag cctctatgga 660

gactcaaatc ctcaaaagtt cacctcatct gctaatggag taaccacaca ttatgtttct 720

cagattggcg acttcccaga tcaaacagaa gacggaggac taccacaaag cggcagaatt 780

gttgttgatt acatgatgca aaaacctggg aaaacaggaa caattgtcta tcaaagaggt 840

gttttgttgc ctcaaaaggt gtggtgcgcg agtggcagga gcaaagtaat aaaagggtca 900

ttgcctttaa ttggtgaagc agattgcctt catgaaaaat acggtggatt aaacaaaagc 960

aagccttact acacaggaga acatgcaaaa gccataggaa attgcccaat atgggtaaaa 1020

acacctttga agcttgccaa tggaaccaaa tatagacctc ctggtggagg atgggaagga 1080

atgattgcag gttggcacgg atacacatct cacggagcac atggagtggc agtggcggca 1140

gaccttaaga gtacacaaga agctataaat aagataacaa aaaatctcaa ttctttgagt 1200

gagctagaag taaagaacct tcaaagacta agtggtgcca tggatgaact ccacaacgaa 1260

atactcgagc tggatgagaa agtggatgat ctcagagctg acactataag ctcacaaata 1320

gaacttgcag tcttgctttc caacgaagga ataataaaca gtgaagacga gcatctattg 1380

gcacttgaga gaaaactaaa gaaaatgctg ggtccctctg ctgtagacat aggaaacgga 1440

tgcttcgaaa ccaaacacaa atgcaaccag acctgcttag acaggatagc tgctggcacc 1500

tttaatgcag gagaattttc tctccccact tttgattcat tgaacattac tgctgcatct 1560

ttaaatgatg atggattgga taaccatact atactgctct attactcaac tgctgcttct 1620

agtttggctg taacattaat gctagctatt tttattgttt atatggtctc cagagacaac 1680

gtttcatgct ccatctgtct ataa 1704

<210> 102
<211> 567
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của HA B Wisconsin (PrL-)

<400> 102

Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
1 5 10 15

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
20 25 30

Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
35 40 45

Thr Thr Pro Thr Lys Ser Tyr Phe Ala Asn Leu Lys Gly Thr Arg Thr

50 55 60

Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp Val
65 70 75 80

Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys Ala
85 90 95

Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile
100 105 110

Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly
115 120 125

Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu Lys
130 135 140

Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn
145 150 155 160

Ala Thr Ser Lys Ile Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro
165 170 175

Lys Asp Asn Tyr Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val Pro
180 185 190

Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His
195 200 205

Ser Asp Asn Lys Thr Gln Met Lys Ser Leu Tyr Gly Asp Ser Asn Pro
210 215 220

Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
225 230 235 240

Gln Ile Gly Asp Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln
245 250 255

Ser Gly Arg Ile Val Val Asp Tyr Met Met Gln Lys Pro Gly Lys Thr
260 265 270

```
Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp
        275                 280                 285

Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile
    290                 295                 300

Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
305                 310                 315                 320

Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
                325                 330                 335

Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
            340                 345                 350

Pro Pro Gly Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr
        355                 360                 365

Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys Ser
    370                 375                 380

Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser
385                 390                 395                 400

Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu
                405                 410                 415

Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg
            420                 425                 430

Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn
        435                 440                 445

Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg
    450                 455                 460

Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Asp Ile Gly Asn Gly
465                 470                 475                 480

Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile
```

485 490 495

Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp
500 505 510

Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn
515 520 525

His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val
530 535 540

Thr Leu Met Leu Ala Ile Phe Ile Val Tyr Met Val Ser Arg Asp Asn
545 550 555 560

Val Ser Cys Ser Ile Cys Leu
565

<210> 103
<211> 1704
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của HA B Wisconsin (PrL-)+H1 California TMCT

<400> 103
atgaaggcaa taattgtact actcatggta gtaacatcca atgcagatcg aatctgcact 60

gggataacat cttcaaactc acctcatgtg gtcaaaacag ctactcaagg ggaggtcaat 120

gtgactggcg tgataccact gacaacaaca ccaacaaaat cttattttgc aaatctcaaa 180

ggaacaagga ccagagggaa actatgcccg gactgtctca actgtacaga tctggatgtg 240

gccttgggca ggccaatgtg tgtggggacc acaccttctg ctaaagcttc aatactccac 300

gaggtcagac ctgttacatc cgggtgcttt cctataatgc acgacagaac aaaaatcagg 360

caactaccca atcttctcag aggatatgaa aatatcaggt tatcaaccca aaacgttatc 420

gatgcagaaa aagcaccagg aggaccctac agacttggaa cctcaggatc ttgccctaac 480

gctaccagta aaatcggatt ttttgcaaca atggcttggg ctgtcccaaa ggacaactac 540

aaaaatgcaa cgaacccact aacagtagaa gtaccataca tttgtacaga aggggaagac 600

caaattactg tttgggggtt ccattcagat aacaaaaccc aaatgaagag cctctatgga 660

gactcaaatc ctcaaaagtt cacctcatct gctaatggag taaccacaca ttatgtttct 720

cagattggcg acttcccaga tcaaacagaa gacggaggac taccacaaag cggcagaatt 780

gttgttgatt acatgatgca aaaacctggg aaaacaggaa caattgtcta tcaaagaggt 840

gttttgttgc ctcaaaaggt gtggtgcgcg agtggcagga gcaaagtaat aaaagggtca 900

ttgcctttaa ttggtgaagc agattgcctt catgaaaaat acggtggatt aaacaaaagc 960

aagccttact acacaggaga acatgcaaaa gccataggaa attgcccaat atgggtaaaa 1020

acacctttga agcttgccaa tggaaccaaa tatagacctc ctggtggagg atgggaagga 1080

atgattgcag gttggcacgg atacacatct cacggagcac atggagtggc agtggcggca 1140

gaccttaaga gtacacaaga agctataaat aagataacaa aaaatctcaa ttctttgagt 1200

gagctagaag taaagaacct tcaaagacta agtggtgcca tggatgaact ccacaacgaa 1260

atactcgagc tggatgagaa agtggatgat ctcagagctg acactataag ctcacaaata 1320

gaacttgcag tcttgctttc caacgaagga ataataaaca gtgaagacga gcatctattg 1380

gcacttgaga gaaaactaaa gaaaatgctg ggtccctctg ctgtagacat aggaaacgga 1440

tgcttcgaaa ccaaacacaa atgcaaccag acctgcttag acaggatagc tgctggcacc 1500

tttaatgcag gagaattttc tctccccact tttgattcat tgaacattac tgctgcatct 1560

ttaaatgatg atggattgga taactaccag attttggcga tctattcaac tgtcgccagt 1620

tcattggtac tggtagtctc cctgggggca atcagtttct ggatgtgctc taatgggtct 1680

ctacagtgta gaatatgtat ttaa 1704

<210> 104
<211> 567
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của HA B Wisconsin (PrL-)+H1 California TMC

<400> 104

Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
1 5 10 15

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
20 25 30

```
Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
        35                  40                  45

Thr Thr Pro Thr Lys Ser Tyr Phe Ala Asn Leu Lys Gly Thr Arg Thr
    50                  55                  60

Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp Val
65                  70                  75                  80

Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys Ala
                85                  90                  95

Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile
            100                 105                 110

Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly
        115                 120                 125

Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu Lys
    130                 135                 140

Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn
145                 150                 155                 160

Ala Thr Ser Lys Ile Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro
                165                 170                 175

Lys Asp Asn Tyr Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val Pro
            180                 185                 190

Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His
        195                 200                 205

Ser Asp Asn Lys Thr Gln Met Lys Ser Leu Tyr Gly Asp Ser Asn Pro
    210                 215                 220

Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
225                 230                 235                 240

Gln Ile Gly Asp Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln
```

245 250 255

Ser Gly Arg Ile Val Val Asp Tyr Met Met Gln Lys Pro Gly Lys Thr
260 265 270

Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp
275 280 285

Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile
290 295 300

Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
305 310 315 320

Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
325 330 335

Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
340 345 350

Pro Pro Gly Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr
355 360 365

Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys Ser
370 375 380

Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser
385 390 395 400

Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu
405 410 415

Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg
420 425 430

Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn
435 440 445

Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg
450 455 460

Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Asp Ile Gly Asn Gly
465 470 475 480

Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile
485 490 495

Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp
500 505 510

Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn
515 520 525

Tyr Gln Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu
530 535 540

Val Val Ser Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser
545 550 555 560

Leu Gln Cys Arg Ile Cys Ile
565

<210> 105
<211> 1734
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Brisbane (PrL-)+H1 California TMCT

<400> 105
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg ccgatcgaat ctgcactgga ataacatcgt caaactcacc acatgtcgtc 120

aaaactgcta ctcaagggga ggtcaatgtg actggtgtaa taccactgac aacaacaccc 180

accaaatctc attttgcaaa tctcaaagga acagaaacca gggggaaact atgcccaaaa 240

tgcctcaact gcacagatct ggacgtagcc ttgggcagac caaaatgcac ggggaaaata 300

ccctcggcaa gagtttcaat actccatgaa gtcagacctg ttacatctgg gtgctttcct 360

ataatgcacg acagaacaaa aattagacag ctgcctaacc ttctccgagg atacgaacat 420

atcaggttat caacccataa cgttatcaat gcagaaaatg caccaggagg accctacaaa 480

attggaacct cagggtcttg ccctaacatt accaatggaa acggattttt cgcaacaatg 540

gcttgggccg tcccaaaaaa cgacaaaaac aaaacagcaa caaatccatt aacaatagaa 600

gtaccataca tttgtacaga aggagaagac caaattaccg tttgggggtt ccactctgac 660

aacgagaccc aaatggcaaa gctctatggg gactcaaagc cccagaagtt cacctcatct 720

gccaacggag tgaccacaca ttacgtttca cagattggtg gcttcccaaa tcaaacagaa 780

gacggaggac taccacaaag tggtagaatt gttgttgatt acatggtgca aaaatctggg 840

aaaacaggaa caattaccta tcaaaggggt attttattgc ctcaaaaggt gtggtgcgca 900

agtggcagga gcaaggtaat aaaaggatcc ttgcctttaa ttggagaagc agattgcctc 960

cacgaaaaat acggtggatt aaacaaaagc aagccttact acacagggga acatgcaaag 1020

gccataggaa attgcccaat atgggtgaaa acacccttga agctggccaa tggaatcaaa 1080

tatagacctc ctggtggagg atgggaagga atgattgcag gttggcacgg atacacatcc 1140

catggggcac atggagtagc ggtggcagca gaccttaaga gcactcaaga ggccataaac 1200

aagataacaa aaaatctcaa ctctttgagt gagctggaag taaagaatct tcaaagacta 1260

agcggtgcca tggatgaact ccacaacgaa atactagaac tagatgagaa agtggatgat 1320

ctcagagctg atacaataag ctcacaaata gaactcgcag tcctgctttc caatgaagga 1380

ataataaaca gtgaagatga acatctcttg gcgcttgaaa gaaagctgaa gaaaatgctg 1440

ggcccctctg ctgtagagat agggaatgga tgctttgaaa ccaaacacaa gtgcaaccag 1500

acctgtctcg acagaatagc tgctggtacc tttgatgcag gagaattttc tctccccacc 1560

tttgattcac tgaatattac tgctgcatct ttaaatgacg atggattgga taattaccag 1620

attttggcga tctattcaac tgtcgccagt tcattggtac tggtagtctc cctgggggca 1680

atcagtttct ggatgtgctc taatgggtct ctacagtgta gaatatgtat ttaa 1734

<210> 106
<211> 577
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Brisbane (PrL-)+H1 California TMCT

<400> 106

```
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                  10                  15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
            20                  25                  30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
        35                  40                  45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
    50                  55                  60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
65                  70                  75                  80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
                85                  90                  95

Thr Gly Lys Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
            100                 105                 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
        115                 120                 125

Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
    130                 135                 140

Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
145                 150                 155                 160

Ile Gly Thr Ser Gly Ser Cys Pro Asn Ile Thr Asn Gly Asn Gly Phe
                165                 170                 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Lys Asn Lys Thr
            180                 185                 190

Ala Thr Asn Pro Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly
        195                 200                 205

Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu Thr Gln
```

210 215 220

Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser
225 230 235 240

Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro
245 250 255

Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val
260 265 270

Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln
275 280 285

Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser
290 295 300

Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu
305 310 315 320

His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly
325 330 335

Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro
340 345 350

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp
355 360 365

Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His
370 375 380

Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn
385 390 395 400

Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn
405 410 415

Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu
420 425 430

Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser
435 440 445

Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser
450 455 460

Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu
465 470 475 480

Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys His
485 490 495

Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asp
500 505 510

Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala
515 520 525

Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn Tyr Gln Ile Leu Ala Ile
530 535 540

Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Val Val Ser Leu Gly Ala
545 550 555 560

Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys
565 570 575

Ile

<210> 107
<211> 1731
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-)

<400> 107
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg ccgatcgaat ctgcactggg ataacatctt caaactcacc tcatgtggtc 120

aaaacagcta ctcaagggga ggtcaatgtg actggtgtga taccactaac aacaacacca 180

acaaaatctt attttgcaaa tctcaaagga acaaagacca gagggaaact atgcccagac 240

tgtctcaact gtacagatct ggatgtggcc ctgggcaggc caatgtgtgt gggaactaca 300

ccttctgcga aagcttcaat acttcacgaa gtcagacctg ttacatccgg gtgcttccct 360

ataatgcacg acagaacaaa aatcaggcaa ctagccaatc ttctcagagg atatgaaaat 420

atcaggttat caacccaaaa cgttatcgat gcagaaaagg caccaggagg accctacaga 480

cttggaacct caggatcttg ccctaacgct accagtaaaa gcggattttt cgcaacaatg 540

gcttgggctg tcccaaagga caacaacaaa aatgcaacga acccattaac agtagaagta 600

ccatacattt gtgcagaagg ggaagaccaa attactgttt gggggttcca ttcagataac 660

aaaacccaaa tgaagaacct ctatggagac tcaaatcctc aaaagttcac ctcatctgct 720

aatggagtaa ccacacatta tgtttctcag attggcggct tcccagatca aacagaagac 780

ggaggactac cacaaagcgg cagaattgtc gttgattaca tgatgcaaaa acctgggaaa 840

acaggaacaa ttgtctatca aagaggtgtt ttgttgcctc aaaaggtgtg gtgcgcgagt 900

ggcaggagca aagtaataaa agggtccttg cctttaattg gtgaagcaga ttgccttcat 960

gaaaaatacg gtggattaaa caaaagcaag ccttactaca caggagaaca tgcaaaagcc 1020

ataggaaatt gcccaatatg ggtgaaaaca cctttgaagc ttgccaatgg aaccaaatat 1080

agacctcctg gtggaggatg ggaaggaatg attgcaggtt ggcacggata cacatctcac 1140

ggagcacatg gagtggcagt tgctgcagac cttaagagca cacaagaagc tataaacaag 1200

ataacaaaaa atctcaactc tttgagtgag ctagaagtaa agaatcttca aaggctaagt 1260

ggtgccatgg atgaactcca caacgaaata ctcgagctgg atgagaaagt ggatgacctc 1320

agagctgaca ctataagttc acaaatagaa cttgcagtct tgctttccaa cgaaggaata 1380

ataaacagtg aagacgagca tctattggca cttgagagaa aactaaagaa aatgctgggt 1440

ccctctgctg tagacatagg aaatggatgc ttcgaaacca aacacaaatg caaccagacc 1500

tgcttagaca ggatagctgc tggcaccttt aatgcaggag agttttctct ccccactttt 1560

gattcattga acattactgc tgcatcttta aatgatgatg gattggataa ccatactata 1620

ctgctctatt actcaactgc tgcttctagt ttggctgtaa cattgatgct agctattttt 1680

attgtttata tggtctccag agacaacgtt tcatgctcca tctgtctata a 1731

<210> 108
<211> 576
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Massachussetts (PrL-)

<400> 108

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser Tyr
50 55 60

Phe Ala Asn Leu Lys Gly Thr Lys Thr Arg Gly Lys Leu Cys Pro Asp
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Met Cys
85 90 95

Val Gly Thr Thr Pro Ser Ala Lys Ala Ser Ile Leu His Glu Val Arg
100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
115 120 125

Arg Gln Leu Ala Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser
130 135 140

Thr Gln Asn Val Ile Asp Ala Glu Lys Ala Pro Gly Gly Pro Tyr Arg
145 150 155 160

Leu Gly Thr Ser Gly Ser Cys Pro Asn Ala Thr Ser Lys Ser Gly Phe
165 170 175

```
Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asp Asn Asn Lys Asn Ala
            180                 185                 190

Thr Asn Pro Leu Thr Val Glu Val Pro Tyr Ile Cys Ala Glu Gly Glu
        195                 200                 205

Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Lys Thr Gln Met
    210                 215                 220

Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala
225                 230                 235                 240

Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro Asp
                245                 250                 255

Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp
            260                 265                 270

Tyr Met Met Gln Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg
        275                 280                 285

Gly Val Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys
    290                 295                 300

Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu His
305                 310                 315                 320

Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly Glu
                325                 330                 335

His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro Leu
            340                 345                 350

Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp Glu
        355                 360                 365

Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His Gly
    370                 375                 380

Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn Lys
```

385 390 395 400

Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn Leu
405 410 415

Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu Glu
420 425 430

Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser Gln
435 440 445

Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser Glu
450 455 460

Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu Gly
465 470 475 480

Pro Ser Ala Val Asp Ile Gly Asn Gly Cys Phe Glu Thr Lys His Lys
485 490 495

Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asn Ala
500 505 510

Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala Ala
515 520 525

Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu Tyr Tyr
530 535 540

Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Leu Ala Ile Phe
545 550 555 560

Ile Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile Cys Leu
565 570 575

<210> 109
<211> 1731
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-)+H1

California TMCT

<400> 109

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg ccgatcgaat ctgcactggg ataacatctt caaactcacc tcatgtggtc 120

aaaacagcta ctcaagggga ggtcaatgtg actggtgtga taccactaac aacaacacca 180

acaaaatctt attttgcaaa tctcaaagga acaaagacca gagggaaact atgcccagac 240

tgtctcaact gtacagatct ggatgtggcc ctgggcaggc caatgtgtgt gggaactaca 300

ccttctgcga aagcttcaat acttcacgaa gtcagacctg ttacatccgg gtgcttccct 360

ataatgcacg acagaacaaa aatcaggcaa ctagccaatc ttctcagagg atatgaaaat 420

atcaggttat caacccaaaa cgttatcgat gcagaaaagg caccaggagg accctacaga 480

cttggaacct caggatcttg ccctaacgct accagtaaaa gcggattttt cgcaacaatg 540

gcttgggctg tcccaaagga caacaacaaa aatgcaacga acccattaac agtagaagta 600

ccatacattt gtgcagaagg ggaagaccaa attactgttt gggggttcca ttcagataac 660

aaaacccaaa tgaagaacct ctatggagac tcaaatcctc aaaagttcac ctcatctgct 720

aatggagtaa ccacacatta tgtttctcag attggcggct tcccagatca aacagaagac 780

ggaggactac cacaaagcgg cagaattgtc gttgattaca tgatgcaaaa acctgggaaa 840

acaggaacaa ttgtctatca aagaggtgtt ttgttgcctc aaaaggtgtg gtgcgcgagt 900

ggcaggagca aagtaataaa agggtccttg cctttaattg gtgaagcaga ttgccttcat 960

gaaaaatacg gtggattaaa caaaagcaag ccttactaca caggagaaca tgcaaaagcc 1020

ataggaaatt gcccaatatg ggtgaaaaca cctttgaagc ttgccaatgg aaccaaatat 1080

agacctcctg gtggaggatg ggaaggaatg attgcaggtt ggcacggata cacatctcac 1140

ggagcacatg gagtggcagt tgctgcagac cttaagagca cacaagaagc tataaacaag 1200

ataacaaaaa atctcaactc tttgagtgag ctagaagtaa agaatcttca aaggctaagt 1260

ggtgccatgg atgaactcca caacgaaata ctcgagctgg atgagaaagt ggatgacctc 1320

agagctgaca ctataagttc acaaatagaa cttgcagtct tgctttccaa cgaaggaata 1380

ataaacagtg aagacgagca tctattggca cttgagagaa aactaaagaa aatgctgggt 1440

ccctctgctg tagacatagg aaatggatgc ttcgaaacca aacacaaatg caaccagacc 1500

tgcttagaca ggatagctgc tggcaccttt aatgcaggag agttttctct ccccactttt 1560

gattcattga acattactgc tgcatcttta aatgatgatg gattggataa ctaccagatt 1620

ttggcgatct attcaactgt cgccagttca ttggtactgg tagtctccct gggggcaatc 1680

agtttctgga tgtgctctaa tgggtctcta cagtgtagaa tatgtattta a 1731

<210> 110
<211> 576
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT

<400> 110

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser Tyr
50 55 60

Phe Ala Asn Leu Lys Gly Thr Lys Thr Arg Gly Lys Leu Cys Pro Asp
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Met Cys
85 90 95

Val Gly Thr Thr Pro Ser Ala Lys Ala Ser Ile Leu His Glu Val Arg
100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
115 120 125

Arg Gln Leu Ala Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser
130 135 140

```
Thr Gln Asn Val Ile Asp Ala Glu Lys Ala Pro Gly Gly Pro Tyr Arg
145                 150                 155                 160

Leu Gly Thr Ser Gly Ser Cys Pro Asn Ala Thr Ser Lys Ser Gly Phe
                165                 170                 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asp Asn Asn Lys Asn Ala
            180                 185                 190

Thr Asn Pro Leu Thr Val Glu Val Pro Tyr Ile Cys Ala Glu Gly Glu
        195                 200                 205             .

Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Lys Thr Gln Met
    210                 215                 220

Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala
225                 230                 235                 240

Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro Asp
                245                 250                 255

Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp
            260                 265                 270

Tyr Met Met Gln Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg
        275                 280                 285

Gly Val Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys
    290                 295                 300

Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu His
305                 310                 315                 320

Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly Glu
                325                 330                 335

His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro Leu
            340                 345                 350

Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp Glu
```

355 360 365

Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His Gly
370 375 380

Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn Lys
385 390 395 400

Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn Leu
405 410 415

Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu Glu
420 425 430

Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser Gln
435 440 445

Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser Glu
450 455 460

Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu Gly
465 470 475 480

Pro Ser Ala Val Asp Ile Gly Asn Gly Cys Phe Glu Thr Lys His Lys
485 490 495

Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asn Ala
500 505 510

Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala Ala
515 520 525

Ser Leu Asn Asp Asp Gly Leu Asp Asn Tyr Gln Ile Leu Ala Ile Tyr
530 535 540

Ser Thr Val Ala Ser Ser Leu Val Leu Val Val Ser Leu Gly Ala Ile
545 550 555 560

Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
565 570 575

<210> 111
<211> 1704
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của HA B Wisconsin (PrL-)

<400> 111
atgaaggcaa taattgtact actcatggta gtaacatcca atgcagatcg aatctgcact 60

gggataacat cttcaaactc acctcatgtg gtcaaaacag ctactcaagg ggaggtcaat 120

gtgactggcg tgataccact gacaacaaca ccaacaaaat cttattttgc aaatctcaaa 180

ggaacaagga ccagagggaa actatgcccg gactgtctca actgtacaga tctggatgtg 240

gccttgggca ggccaatgtg tgtggggacc acaccttctg ctaaagcttc aatactccac 300

gaggtcagac ctgttacatc cgggtgcttt cctataatgc acgacagaac aaaaatcagg 360

caactaccca atcttctcag aggatatgaa aatatcaggt tatcaaccca aaacgttatc 420

gatgcagaaa aagcaccagg aggaccctac agacttggaa cctcaggatc ttgccctaac 480

gctaccagta aaatcggatt ttttgcaaca atggcttggg ctgtcccaaa ggacaactac 540

aaaaatgcaa cgaacccact aacagtagaa gtaccataca tttgtacaga aggggaagac 600

caaattactg tttgggggtt ccattcagat aacaaaaccc aaatgaagag cctctatgga 660

gactcaaatc ctcaaaagtt cacctcatct gctaatggag taaccacaca ttatgtttct 720

cagattggcg acttcccaga tcaaacagaa gacggaggac taccacaaag cggcagaatt 780

gttgttgatt acatgatgca aaaacctggg aaaacaggaa caattgtcta tcaaagaggt 840

gttttgttgc ctcaaaaggt gtggtgcgcg agtggcagga gcaaagtaat aaaagggtca 900

ttgcctttaa ttggtgaagc agattgcctt catgaaaaat acggtggatt aaacaaaagc 960

aagccttact acacaggaga acatgcaaaa gccataggaa attgcccaat atgggtaaaa 1020

acacctttga agcttgccaa tggaaccaaa tatagacctc ctggtggagg atgggaagga 1080

atgattgcag gttggcacgg atacacatct cacggagcac atggagtggc agtggcggca 1140

gaccttaaga gtacacaaga agctataaat aagataacaa aaaatctcaa ttctttgagt 1200

gagctagaag taaagaacct tcaaagacta agtggtgcca tggatgaact ccacaacgaa 1260

atactcgagc tggatgagaa agtggatgat ctcagagctg acactataag ctcacaaata 1320

gaacttgcag tcttgctttc caacgaagga ataataaaca gtgaagacga gcatctattg 1380

gcacttgaga gaaaactaaa gaaaatgctg ggtccctctg ctgtagacat aggaaacgga 1440

tgcttcgaaa ccaaacacaa atgcaaccag acctgcttag acaggatagc tgctggcacc 1500

tttaatgcag gagaattttc tctccccact tttgattcat tgaacattac tgctgcatct 1560

ttaaatgatg atggattgga taaccatact atactgctct attactcaac tgctgcttct 1620

agtttggctg taacattaat gctagctatt tttattgttt atatggtctc cagagacaac 1680

gtttcatgct ccatctgtct ataa 1704

<210> 112
<211> 567
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của HA B Wisconsin (PrL-)

<400> 112

Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
1 5 10 15

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
20 25 30

Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
35 40 45

Thr Thr Pro Thr Lys Ser Tyr Phe Ala Asn Leu Lys Gly Thr Arg Thr
50 55 60

Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp Val
65 70 75 80

Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys Ala
85 90 95

Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile
100 105 110

Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly

```
        115                 120                 125

Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu Lys
    130                 135                 140

Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn
145                 150                 155                 160

Ala Thr Ser Lys Ile Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro
                165                 170                 175

Lys Asp Asn Tyr Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val Pro
            180                 185                 190

Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His
        195                 200                 205

Ser Asp Asn Lys Thr Gln Met Lys Ser Leu Tyr Gly Asp Ser Asn Pro
    210                 215                 220

Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
225                 230                 235                 240

Gln Ile Gly Asp Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln
                245                 250                 255

Ser Gly Arg Ile Val Val Asp Tyr Met Met Gln Lys Pro Gly Lys Thr
            260                 265                 270

Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp
        275                 280                 285

Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile
    290                 295                 300

Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser
305                 310                 315                 320

Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
                325                 330                 335
```

```
Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
            340                 345                 350

Pro Pro Gly Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr
        355                 360                 365

Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys Ser
    370                 375                 380

Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser
385                 390                 395                 400

Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu
                405                 410                 415

Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg
            420                 425                 430

Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn
        435                 440                 445

Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg
    450                 455                 460

Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Asp Ile Gly Asn Gly
465                 470                 475                 480

Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile
                485                 490                 495

Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp
            500                 505                 510

Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn
        515                 520                 525

His Thr Ile Leu Leu Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val
    530                 535                 540

Thr Leu Met Leu Ala Ile Phe Ile Val Tyr Met Val Ser Arg Asp Asn
```

545 550 555 560

Val Ser Cys Ser Ile Cys Leu
565

<210> 113
<211> 1704
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của HA B Wisconsin (PrL-)+H1 California TMCT

<400> 113
atgaaggcaa taattgtact actcatggta gtaacatcca atgcagatcg aatctgcact 60

gggataacat cttcaaactc acctcatgtg gtcaaaacag ctactcaagg ggaggtcaat 120

gtgactggcg tgataccact gacaacaaca ccaacaaaat cttattttgc aaatctcaaa 180

ggaacaagga ccagagggaa actatgcccg gactgtctca actgtacaga tctggatgtg 240

gccttgggca ggccaatgtg tgtggggacc acaccttctg ctaaagcttc aatactccac 300

gaggtcagac ctgttacatc cgggtgcttt cctataatgc acgacagaac aaaaatcagg 360

caactaccca atcttctcag aggatatgaa aatatcaggt tatcaaccca aaacgttatc 420

gatgcagaaa aagcaccagg aggaccctac agacttggaa cctcaggatc ttgccctaac 480

gctaccagta aaatcggatt ttttgcaaca atggcttggg ctgtcccaaa ggacaactac 540

aaaaatgcaa cgaacccact aacagtagaa gtaccataca tttgtacaga aggggaagac 600

caaattactg tttgggggtt ccattcagat aacaaaaccc aaatgaagag cctctatgga 660

gactcaaatc ctcaaaagtt cacctcatct gctaatggag taaccacaca ttatgtttct 720

cagattggcg acttcccaga tcaaacagaa gacggaggac taccacaaag cggcagaatt 780

gttgttgatt acatgatgca aaaacctggg aaaacaggaa caattgtcta tcaaagaggt 840

gttttgttgc ctcaaaaggt gtggtgcgcg agtggcagga gcaaagtaat aaaagggtca 900

ttgcctttaa ttggtgaagc agattgcctt catgaaaaat acggtggatt aaacaaaagc 960

aagccttact acacaggaga acatgcaaaa gccataggaa attgcccaat atgggtaaaa 1020

acacctttga agcttgccaa tggaaccaaa tatagacctc ctggtggagg atgggaagga 1080

atgattgcag gttggcacgg atacacatct cacggagcac atggagtggc agtggcggca 1140

gaccttaaga gtacacaaga agctataaat aagataacaa aaaatctcaa ttctttgagt 1200

gagctagaag taaagaacct tcaaagacta agtggtgcca tggatgaact ccacaacgaa 1260

atactcgagc tggatgagaa agtggatgat ctcagagctg acactataag ctcacaaata 1320

gaacttgcag tcttgctttc caacgaagga ataataaaca gtgaagacga gcatctattg 1380

gcacttgaga gaaaactaaa gaaaatgctg ggtccctctg ctgtagacat aggaaacgga 1440

tgcttcgaaa ccaaacacaa atgcaaccag acctgcttag acaggatagc tgctggcacc 1500

tttaatgcag gagaattttc tctccccact tttgattcat tgaacattac tgctgcatct 1560

ttaaatgatg atggattgga taactaccag attttggcga tctattcaac tgtcgccagt 1620

tcattggtac tggtagtctc cctgggggca atcagtttct ggatgtgctc taatgggtct 1680

ctacagtgta gaatatgtat ttaa 1704

<210> 114
<211> 567
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của HA B Wisconsin (PrL-)+H1 California TMC.

<400> 114

Met Lys Ala Ile Ile Val Leu Leu Met Val Val Thr Ser Asn Ala Asp
1 5 10 15

Arg Ile Cys Thr Gly Ile Thr Ser Ser Asn Ser Pro His Val Val Lys
20 25 30

Thr Ala Thr Gln Gly Glu Val Asn Val Thr Gly Val Ile Pro Leu Thr
35 40 45

Thr Thr Pro Thr Lys Ser Tyr Phe Ala Asn Leu Lys Gly Thr Arg Thr
50 55 60

Arg Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys Thr Asp Leu Asp Val
65 70 75 80

Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr Pro Ser Ala Lys Ala
85 90 95

Ser Ile Leu His Glu Val Arg Pro Val Thr Ser Gly Cys Phe Pro Ile
100 105 110

Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro Asn Leu Leu Arg Gly
115 120 125

Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu Lys
130 135 140

Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro Asn
145 150 155 160

Ala Thr Ser Lys Ile Gly Phe Phe Ala Thr Met Ala Trp Ala Val Pro
165 170 175

Lys Asp Asn Tyr Lys Asn Ala Thr Asn Pro Leu Thr Val Glu Val Pro
180 185 190

Tyr Ile Cys Thr Glu Gly Glu Asp Gln Ile Thr Val Trp Gly Phe His
195 200 205

Ser Asp Asn Lys Thr Gln Met Lys Ser Leu Tyr Gly Asp Ser Asn Pro
210 215 220

Gln Lys Phe Thr Ser Ser Ala Asn Gly Val Thr Thr His Tyr Val Ser
225 230 235 240

Gln Ile Gly Asp Phe Pro Asp Gln Thr Glu Asp Gly Gly Leu Pro Gln
245 250 255

Ser Gly Arg Ile Val Val Asp Tyr Met Met Gln Lys Pro Gly Lys Thr
260 265 270

Gly Thr Ile Val Tyr Gln Arg Gly Val Leu Leu Pro Gln Lys Val Trp
275 280 285

Cys Ala Ser Gly Arg Ser Lys Val Ile Lys Gly Ser Leu Pro Leu Ile
290 295 300

Gly Glu Ala Asp Cys Leu His Glu Lys Tyr Gly Gly Leu Asn Lys Ser

305 310 315 320

Lys Pro Tyr Tyr Thr Gly Glu His Ala Lys Ala Ile Gly Asn Cys Pro
325 330 335

Ile Trp Val Lys Thr Pro Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg
340 345 350

Pro Pro Gly Gly Gly Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr
355 360 365

Thr Ser His Gly Ala His Gly Val Ala Val Ala Ala Asp Leu Lys Ser
370 375 380

Thr Gln Glu Ala Ile Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser
385 390 395 400

Glu Leu Glu Val Lys Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu
405 410 415

Leu His Asn Glu Ile Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg
420 425 430

Ala Asp Thr Ile Ser Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn
435 440 445

Glu Gly Ile Ile Asn Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg
450 455 460

Lys Leu Lys Lys Met Leu Gly Pro Ser Ala Val Asp Ile Gly Asn Gly
465 470 475 480

Cys Phe Glu Thr Lys His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile
485 490 495

Ala Ala Gly Thr Phe Asn Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp
500 505 510

Ser Leu Asn Ile Thr Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn
515 520 525

Tyr Gln Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu
530 535 540

Val Val Ser Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser
545 550 555 560

Leu Gln Cys Arg Ile Cys Ile
565

<210> 115
<211> 1707
<212> ADN
<213> Cúm H5 Indonesia

<400> 115
atggagaaaa tagtgcttct tcttgcaata gtcagtcttg ttaaaagtga tcagatttgc 60

attggttacc atgcaaacaa ttcaacagag caggttgaca caatcatgga aaagaacgtt 120

actgttacac atgcccaaga catactggaa aagacacaca acgggaagct ctgcgatcta 180

gatggagtga agcctctaat tttaagagat tgtagtgtag ctggatggct cctcgggaac 240

ccaatgtgtg acgaattcat caatgtaccg gaatggtctt acatagtgga gaaggccaat 300

ccaaccaatg acctctgtta cccagggagt ttcaacgact atgaagaact gaaacaccta 360

ttgagcagaa taaaccattt tgagaaaatt caaatcatcc ccaaaagttc ttggtccgat 420

catgaagcct catcaggagt tagctcagca tgtccatacc tgggaagtcc ctccttttt 480

agaaatgtgg tatggcttat caaaaagaac agtacatacc caacaataaa gaaaagctac 540

aataatacca accaagagga tcttttggta ctgtggggaa ttcaccatcc taatgatgcg 600

gcagagcaga caaggctata tcaaaaccca accacctata tttccattgg gacatcaaca 660

ctaaaccaga gattggtacc aaaaatagct actagatcca aagtaaacgg gcaaagtgga 720

aggatggagt tcttctggac aattttaaaa cctaatgatg caatcaactt cgagagtaat 780

ggaaatttca ttgctccaga atatgcatac aaaattgtca agaaagggga ctcagcaatt 840

atgaaaagtg aattggaata tggtaactgc aacaccaagt gtcaaactcc aatgggggcg 900

ataaactcta gtatgccatt ccacaacata caccctctca ccatcgggga atgccccaaa 960

tatgtgaaat caaacagatt agtccttgca acagggctca gaaatagccc tcaaagagag 1020

agcagaagaa aaaagagagg actatttgga gctatagcag gttttataga gggaggatgg 1080

cagggaatgg tagatggttg gtatgggtac caccatagca atgagcaggg gagtgggtac 1140

gctgcagaca aagaatccac tcaaaaggca atagatggag tcaccaataa ggtcaactca 1200

atcattgaca aaatgaacac tcagtttgag gccgttggaa gggaatttaa taacttagaa 1260

aggagaatag agaatttaaa caagaagatg gaagacgggt ttctagatgt ctggacttat 1320

aatgccgaac ttctggttct catggaaaat gagagaactc tagactttca tgactcaaat 1380

gttaagaacc tctacgacaa ggtccgacta cagcttaggg ataatgcaaa ggagctgggt 1440

aacggttgtt tcgagttcta tcacaaatgt gataatgaat gtatggaaag tataagaaac 1500

ggaacgtaca actatccgca gtattcagaa gaagcaagat taaaaagaga ggaaataagt 1560

ggggtaaaat tggaatcaat aggaacttac caaatactgt caatttattc aacagtggcg 1620

agttccctag cactggcaat catgatggct ggtctatctt tatggatgtg ctccaatgga 1680

tcgttacaat gcagaatttg catttaa 1707

<210> 116
<211> 568
<212> PRT
<213> Cúm H5 Indonesia

<400> 116

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
1 5 10 15

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20 25 30

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
35 40 45

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
50 55 60

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
65 70 75 80

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
85 90 95

```
Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn
            100                 105                 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
        115                 120                 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
    130                 135                 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
145                 150                 155                 160

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
                165                 170                 175

Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
            180                 185                 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
        195                 200                 205

Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
    210                 215                 220

Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
225                 230                 235                 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
                245                 250                 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
            260                 265                 270

Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
        275                 280                 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
    290                 295                 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
```

```
305                 310                 315                 320

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
                325                 330                 335

Pro Gln Arg Glu Ser Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
            340                 345                 350

Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
        355                 360                 365

Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
    370                 375                 380

Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
385                 390                 395                 400

Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
                405                 410                 415

Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
            420                 425                 430

Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
        435                 440                 445

Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
    450                 455                 460

Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
465                 470                 475                 480

Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
                485                 490                 495

Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala
            500                 505                 510

Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
        515                 520                 525
```

Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
530 535 540

Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
545 550 555 560

Ser Leu Gln Cys Arg Ile Cys Ile
565

<210> 117
<211> 52
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1800 IF**(SacII)-PDI.s1+4c

<400> 117
acagggccca ataccgcgga gaaaatggcg aaaaacgttg cgattttcgg ct 52

<210> 118
<211> 51
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1800 IF-H3V36111.s1-4r

<400> 118
actaaagaaa ataggccttc aaatgcaaat gttgcaccta atgttgccct t 51

<210> 119
<211> 1725
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/H3 Victoria.

<400> 119
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cccaaaaact tcctggaaat gacaacagca cggcaacgct gtgccttggg 120

caccatgcag taccaaacgg aacgatagtg aaaacaatca cgaatgacca aattgaagtt 180

actaatgcta ctgagctggt tcagaattcc tcaataggtg aaatatgcga cagtcctcat 240

cagatccttg atggagaaaa ctgcacacta atagatgctc tattgggaga ccctcagtgt 300

gatggcttcc aaaataagaa atgggacctt tttgttgaac gaagcaaagc ctacagcaac 360

tgttaccctt atgatgtgcc ggattatgcc tcccttaggt cactagttgc ctcatccggc 420

acactggagt ttaacaatga aagcttcaat tggactggag tcactcaaaa cggaacaagt 480

tctgcttgca taaggagatc taataatagt ttctttagta gattaaattg gttgacccac 540

ttaaacttca aatacccagc attgaacgtg actatgccaa acaatgaaca atttgacaaa 600

ttgtacattt ggggggttca ccacccgggt acggacaagg accaaatctt cctgtatgct 660

caatcatcag gaagaatcac agtatctacc aaaagaagcc aacaagctgt aatcccgaat 720

atcggatcta gacccagaat aaggaatatc cctagcagaa taagcatcta ttggacaata 780

gtaaaaccgg gagacatact tttgattaac agcacaggga atctaattgc tcctaggggt 840

tacttcaaaa tacgaagtgg gaaaagctca ataatgagat cagatgcacc cattggcaaa 900

tgcaattctg aatgcatcac tccaaatgga agcattccca atgacaaacc attccaaaat 960

gtaaacagga tcacatacgg ggcctgtccc agatatgtta agcaaagcac tctgaaattg 1020

gcaacaggaa tgcgaaatgt accagagaaa caaactagag gcatatttgg cgcaatagcg 1080

ggtttcatag aaaatggttg ggagggaatg gtggatggtt ggtacggttt caggcatcaa 1140

aattctgagg gaagaggaca agcagcagat ctcaaaagca ctcaagcagc aatcgatcaa 1200

atcaatggga agctgaatcg attgatcggg aaaaccaacg agaaattcca tcagattgaa 1260

aaagaattct cagaagtcga agggagaatt caggaccttg agaaatatgt tgaggacact 1320

aaaatagatc tctggtcata caacgcggag cttcttgttg ccctggagaa ccaacataca 1380

attgatctaa ctgactcaga aatgaacaaa ctgtttgaaa aaacaaagaa gcaactaagg 1440

gaaaatgctg aggatatggg caatggttgt ttcaaaatat accacaaatg tgacaatgcc 1500

tgcataggat caatcagaaa tggaacttat gaccacgatg tatacagaga tgaagcatta 1560

aacaaccggt tccagatcaa gggagttgag ctgaagtcag ggtacaaaga ttggatccta 1620

tggatttcct ttgccatatc atgtttttg ctttgtgttg ctttgttggg gttcatcatg 1680

tgggcctgcc aaaagggcaa cattaggtgc aacatttgca tttga 1725

<210> 120
<211> 4644
<212> ADN

<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 2171

<400> 120
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120

aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180

aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240

ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300

gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360

aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420

aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480

taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga 540

aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600

atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660

taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcattttta 720

tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg 780

gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840

acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat 900

ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960

accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt 1020

gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080

aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg 1140

gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg 1200

actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc 1260

aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg 1320

gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca 1380

tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt 1440

agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg 1500

tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga 1560

tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt 1620

ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa 1680

tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac 1740

ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg 1800

cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa 1860

gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt 1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct 1980

ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc 2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg 2100

cgcgttggga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tggtggagca 2160

cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat 2220

tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gcccagctat 2280

ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat gccatcattg 2340

cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtggtccca aagatggacc 2400

cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt 2460

ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga 2520

tacagtctca gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa 2580

cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga 2640

aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc 2700

tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga 2760

cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga 2820

tgacgcacaa tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca 2880

tttggagagg tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa 2940

ccaaaccttc ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc 3000

ttgcgtgagc gatcttcaac gttgtcagat cgtgcttcgg caccagtaca gggcccaata 3060

ccgcggagaa aatggcgaaa aacgttgcga ttttcggctt attgttttct cttcttgtgt 3120

tggttccttc tcagatcttc gcgacgtcac tcctcagcca aaacgacacc cccatctgtc 3180

tatccactgg cccctggatc tgctgcccaa actaactcca tggtgaccct gggatgcctg 3240

gtcaagggct atttccctga gccagtgaca gtgacctgga actctggatc cctgtccagc 3300

ggtgtgcaca ccttcccagc tgtcctgcag tctgacctct acactctgag cagctcagtg 3360

actgtcccct ccagcacctg gcccagcgag accgtcacct gcaacgttgc ccacccggcc 3420

agcagcacca aggtggacaa gaaaattgtg cccagggatt gtggttgtaa gccttgcata 3480

tgtacagtcc cagaagtatc atctgtcttc atcttccccc caaagcccaa ggatgtgctc 3540

accattactc tgactcctaa ggtcacgtgt gttgtggtag acatcagcaa ggatgatccc 3600

gaggtccagt tcagctggtt tgtagatgat gtggaggtgc acacagctca gacgcaaccc 3660

cgggaggagc agttcaacag cactttccgc tcagtcagtg aacttcccat catgcaccag 3720

gactggctca atggcaagga gacgtccaga ttttggcgat ctattcaact gtcgccagtt 3780

cattggtact ggtagtctcc ctgggggcaa tcagtttctg gatgtgctct aatgggtctc 3840

tacagtgtag aatatgtatt taaaggccta ttttctttag tttgaattta ctgttattcg 3900

gtgtgcattt ctatgtttgg tgagcggttt tctgtgctca gagtgtgttt attttatgta 3960

atttaatttc tttgtgagct cctgtttagc aggtcgtccc ttcagcaagg acacaaaaag 4020

attttaattt tattaaaaaa aaaaaaaaaa aagaccggga attcgatatc aagcttatcg 4080

acctgcagat cgttcaaaca tttggcaata aagtttctta agattgaatc ctgttgccgg 4140

tcttgcgatg attatcatat aatttctgtt gaattacgtt aagcatgtaa taattaacat 4200

gtaatgcatg acgttattta tgagatgggt ttttatgatt agagtcccgc aattatacat 4260

ttaatacgcg atagaaaaca aaatatagcg cgcaaactag gataaattat cgcgcgcggt 4320

gtcatctatg ttactagatc tctagagtct caagcttggc gcgcccacgt gactagtggc 4380

actggccgtc gttttacaac gtcgtgactg ggaaaaccct ggcgttaccc aacttaatcg 4440

ccttgcagca catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg 4500

cccttcccaa cagttgcgca gcctgaatgg cgaatgctag agcagcttga gcttggatca 4560

gattgtcgtt tcccgccttc agtttaaact atcagtgttt gacaggatat attggcgggt 4620

aaacctaaga gaaaagagcg ttta 4644

<210> 121
<211> 3129
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1800

<400> 121
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaggg cccaataccg cggagaaaat ggcgaaaaac gttgcgattt tcggcttatt 960

gttttctctt cttgtgttgg ttccttctca gatcttcgcc caaaaacttc ctggaaatga 1020

caacagcacg gcaacgctgt gccttgggca ccatgcagta ccaaacggaa cgatagtgaa 1080

aacaatcacg aatgaccaaa ttgaagttac taatgctact gagctggttc agaattcctc 1140

aataggtgaa atatgcgaca gtcctcatca gatccttgat ggagaaaact gcacactaat 1200

agatgctcta ttgggagacc ctcagtgtga tggcttccaa aataagaaat gggacctttt 1260

tgttgaacga agcaaagcct acagcaactg ttacccttat gatgtgccgg attatgcctc 1320

ccttaggtca ctagttgcct catccggcac actggagttt aacaatgaaa gcttcaattg 1380

gactggagtc actcaaaacg gaacaagttc tgcttgcata aggagatcta ataatagttt 1440

ctttagtaga ttaaattggt tgacccactt aaacttcaaa tacccagcat tgaacgtgac 1500

tatgccaaac aatgaacaat ttgacaaatt gtacatttgg ggggttcacc acccgggtac 1560

ggacaaggac caaatcttcc tgtatgctca atcatcagga agaatcacag tatctaccaa 1620

aagaagccaa caagctgtaa tcccgaatat cggatctaga cccagaataa ggaatatccc 1680

tagcagaata agcatctatt ggacaatagt aaaaccggga gacatacttt tgattaacag 1740

cacagggaat ctaattgctc ctaggggtta cttcaaaata cgaagtggga aaagctcaat 1800

aatgagatca gatgcaccca ttggcaaatg caattctgaa tgcatcactc caaatggaag 1860

cattcccaat gacaaaccat tccaaaatgt aaacaggatc acatacgggg cctgtcccag 1920

atatgttaag caaagcactc tgaaattggc aacaggaatg cgaaatgtac cagagaaaca 1980

aactagaggc atatttggcg caatagcggg tttcatagaa aatggttggg agggaatggt 2040

ggatggttgg tacggtttca ggcatcaaaa ttctgaggga agaggacaag cagcagatct 2100

caaaagcact caagcagcaa tcgatcaaat caatgggaag ctgaatcgat tgatcgggaa 2160

aaccaacgag aaattccatc agattgaaaa agaattctca gaagtcgaag ggagaattca 2220

ggaccttgag aaatatgttg aggacactaa aatagatctc tggtcataca acgcggagct 2280

tcttgttgcc ctggagaacc aacatacaat tgatctaact gactcagaaa tgaacaaact 2340

gtttgaaaaa acaaagaagc aactaaggga aaatgctgag gatatgggca atggttgttt 2400

caaaatatac cacaaatgtg acaatgcctg cataggatca atcagaaatg gaacttatga 2460

ccacgatgta tacagagatg aagcattaaa caaccggttc cagatcaagg gagttgagct 2520

gaagtcaggg tacaaagatt ggatcctatg gatttccttt gccatatcat gtttttgct 2580

ttgtgttgct ttgttggggt tcatcatgtg ggcctgccaa aagggcaaca ttaggtgcaa 2640

catttgcatt tgaaggccta ttttctttag tttgaattta ctgttattcg gtgtgcattt 2700

ctatgtttgg tgagcggttt tctgtgctca gagtgtgttt attttatgta atttaatttc 2760

tttgtgagct cctgtttagc aggtcgtccc ttcagcaagg acacaaaaag attttaattt 2820

tattaaaaaa aaaaaaaaaa aagaccggga attcgatatc aagcttatcg acctgcagat 2880

cgttcaaaca tttggcaata aagtttctta agattgaatc ctgttgccgg tcttgcgatg 2940

attatcatat aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgcatg 3000

acgttattta tgagatgggt ttttatgatt agagtcccgc aattatacat ttaatacgcg 3060

atagaaaaca aaatatagcg cgcaaactag gataaattat cgcgcgcggt gtcatctatg 3120

ttactagat 3129

<210> 122
<211> 574
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H3 Victoria

<400> 122

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Lys Leu Pro Gly Asn Asp Asn
20 25 30

Ser Thr Ala Thr Leu Cys Leu Gly His His Ala Val Pro Asn Gly Thr
35 40 45

Ile Val Lys Thr Ile Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr
50 55 60

Glu Leu Val Gln Asn Ser Ser Ile Gly Glu Ile Cys Asp Ser Pro His
65 70 75 80

Gln Ile Leu Asp Gly Glu Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly
85 90 95

Asp Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val
100 105 110

Glu Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp
115 120 125

Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe
130 135 140

```
Asn Asn Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser
145                 150                 155                 160

Ser Ala Cys Ile Arg Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn
                165                 170                 175

Trp Leu Thr His Leu Asn Phe Lys Tyr Pro Ala Leu Asn Val Thr Met
            180                 185                 190

Pro Asn Asn Glu Gln Phe Asp Lys Leu Tyr Ile Trp Gly Val His His
        195                 200                 205

Pro Gly Thr Asp Lys Asp Gln Ile Phe Leu Tyr Ala Gln Ser Ser Gly
    210                 215                 220

Arg Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Ala Val Ile Pro Asn
225                 230                 235                 240

Ile Gly Ser Arg Pro Arg Ile Arg Asn Ile Pro Ser Arg Ile Ser Ile
                245                 250                 255

Tyr Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr
            260                 265                 270

Gly Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys
        275                 280                 285

Ser Ser Ile Met Arg Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu
    290                 295                 300

Cys Ile Thr Pro Asn Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn
305                 310                 315                 320

Val Asn Arg Ile Thr Tyr Gly Ala Cys Pro Arg Tyr Val Lys Gln Ser
                325                 330                 335

Thr Leu Lys Leu Ala Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr
            340                 345                 350

Arg Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu
```

355 360 365

Gly Met Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly
370 375 380

Arg Gly Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp Gln
385 390 395 400

Ile Asn Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys Phe
405 410 415

His Gln Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln Asp
420 425 430

Leu Glu Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr Asn
435 440 445

Ala Glu Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr
450 455 460

Asp Ser Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu Arg
465 470 475 480

Glu Asn Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys
485 490 495

Cys Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His
500 505 510

Asp Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly
515 520 525

Val Glu Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp Ile Ser Phe
530 535 540

Ala Ile Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly Phe Ile Met
545 550 555 560

Trp Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
565 570

<210> 123
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1819 IF(SacII)-Kozac_PDI.c

<400> 123
gtcgggccca ataccgcgga gaaaatggcg aaaaacgttg cgattttc 48

<210> 124
<211> 51
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1819 IF-H3V36111.s1-4r

<400> 124
actaaagaaa ataggccttc aaatgcaaat gttgcaccta atgttgccct t 51

<210> 125
<211> 4914
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 2181

<400> 125
tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60

gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120

aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180

aatattcacc tactactgtt ataaatcatt attaaacatt agagtaaaga aatatggatg 240

ataagaacaa gagtagtgat attttgacaa caattttgtt gcaacatttg agaaaatttt 300

gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360

aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaaa atagttgtac 420

aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480

taaggatgac gcattagaga gatgtaccat tagagaattt ttggcaagtc attaaaaaga 540

aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600

atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660

taatttgaca tttgatcttt tcctatatat tgccccatag agtcagttaa ctcatttta 720

tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg 780

gtatatttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840

acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat 900

ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960

accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt 1020

gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080

aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg 1140

gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg 1200

actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc 1260

aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg 1320

gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca 1380

tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt 1440

agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg 1500

tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga 1560

tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt 1620

ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa 1680

tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac 1740

ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg 1800

cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa 1860

gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt 1920

tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct 1980

ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc 2040

ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg 2100

cgcgttggga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tggtggagca 2160

cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat 2220

tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gcccagctat 2280

ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat gccatcattg 2340

cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtggtccca aagatggacc 2400

cccacccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt 2460

ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga 2520

tacagtctca gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa 2580

cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga 2640

aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc 2700

tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga 2760

cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga 2820

tgacgcacaa tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca 2880

tttggagagg tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa 2940

ccaaaccttc ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc 3000

ttgcgtgagc gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgttttctt 3060

tcactgaagc gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg 3120

tgtacttgtc ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct 3180

gttcagcccc atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct 3240

acttctgctt gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt 3300

tctataagaa atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga 3360

gaaagattgt taagcttctg tatattctgc ccaaatttgt cgggcccaat accgcggaga 3420

aaatggcgaa aaacgttgcg attttcggct tattgttttc tcttcttgtg ttggttcctt 3480

ctcagatctt cgcgacgtca ctcctcagcc aaaacgacac ccccatctgt ctatccactg 3540

gcccctggat ctgctgccca aactaactcc atggtgaccc tgggatgcct ggtcaagggc 3600

tatttccctg agccagtgac agtgacctgg aactctggat ccctgtccag cggtgtgcac 3660

accttcccag ctgtcctgca gtctgacctc tacactctga gcagctcagt gactgtcccc 3720

tccagcacct ggcccagcga gaccgtcacc tgcaacgttg cccacccggc cagcagcacc 3780

aaggtggaca agaaaattgt gcccagggat tgtggttgta agccttgcat atgtacagtc 3840

ccagaagtat catctgtctt catcttcccc ccaaagccca aggatgtgct caccattact 3900

ctgactccta aggtcacgtg tgttgtggta gacatcagca aggatgatcc cgaggtccag 3960

ttcagctggt ttgtagatga tgtggaggtg cacacagctc agacgcaacc ccgggaggag 4020

cagttcaaca gcactttccg ctcagtcagt gaacttccca tcatgcacca ggactggctc 4080

aatggcaagg agcgatcgct caccatcacc atcaccatca ccatcaccat taaaggccta 4140

ttttctttag tttgaattta ctgttattcg gtgtgcattt ctatgtttgg tgagcggttt 4200

tctgtgctca gagtgtgttt attttatgta atttaatttc tttgtgagct cctgtttagc 4260

aggtcgtccc ttcagcaagg acacaaaaag attttaattt tattaaaaaa aaaaaaaaaa 4320

aagaccggga attcgatatc aagcttatcg acctgcagat cgttcaaaca tttggcaata 4380

aagtttctta agattgaatc ctgttgccgg tcttgcgatg attatcatat aatttctgtt 4440

gaattacgtt aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt 4500

ttttatgatt agagtcccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg 4560

cgcaaactag gataaattat cgcgcgcggt gtcatctatg ttactagatc tctagagtct 4620

caagcttggc gcgcccacgt gactagtggc actggccgtc gttttacaac gtcgtgactg 4680

ggaaaaccct ggcgttaccc aacttaatcg ccttgcagca catccccctt tcgccagctg 4740

gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca gcctgaatgg 4800

cgaatgctag agcagcttga gcttggatca gattgtcgtt tcccgccttc agtttaaact 4860

atcagtgttt gacaggatat attggcgggt aaacctaaga gaaaagagcg ttta 4914

<210> 126
<211> 3480
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc 1819

<400> 126
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta ataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccaatacc gcggagaaaa tggcgaaaaa cgttgcgatt ttcggcttat tgttttctct 1320

tcttgtgttg gttccttctc agatcttcgc ccaaaaactt cctggaaatg acaacagcac 1380

ggcaacgctg tgccttgggc accatgcagt accaaacgga acgatagtga aaacaatcac 1440

gaatgaccaa attgaagtta ctaatgctac tgagctggtt cagaattcct caataggtga 1500

aatatgcgac agtcctcatc agatccttga tggagaaaac tgcacactaa tagatgctct 1560

attgggagac cctcagtgtg atggcttcca aaataagaaa tgggaccttt ttgttgaacg 1620

aagcaaagcc tacagcaact gttaccctta tgatgtgccg gattatgcct cccttaggtc 1680

actagttgcc tcatccggca cactggagtt taacaatgaa agcttcaatt ggactggagt 1740

cactcaaaac ggaacaagtt ctgcttgcat aaggagatct aataatagtt tctttagtag 1800

attaaattgg ttgacccact taaacttcaa atacccagca ttgaacgtga ctatgccaaa 1860

caatgaacaa tttgacaaat tgtacatttg gggggttcac cacccgggta cggacaagga 1920

ccaaatcttc ctgtatgctc aatcatcagg aagaatcaca gtatctacca aaagaagcca 1980

acaagctgta atcccgaata tcggatctag acccagaata aggaatatcc ctagcagaat 2040

aagcatctat tggacaatag taaaaccggg agacatactt ttgattaaca gcacagggaa 2100

tctaattgct cctaggggtt acttcaaaat acgaagtggg aaaagctcaa taatgagatc 2160

agatgcaccc attggcaaat gcaattctga atgcatcact ccaaatggaa gcattcccaa 2220

tgacaaacca ttccaaaatg taaacaggat cacatacggg gcctgtccca gatatgttaa 2280

gcaaagcact ctgaaattgg caacaggaat gcgaaatgta ccagagaaac aaactagagg 2340

catatttggc gcaatagcgg gtttcataga aaatggttgg gagggaatgg tggatggttg 2400

gtacggtttc aggcatcaaa attctgaggg aagaggacaa gcagcagatc tcaaaagcac 2460

tcaagcagca atcgatcaaa tcaatgggaa gctgaatcga ttgatcggga aaaccaacga 2520

gaaattccat cagattgaaa aagaattctc agaagtcgaa gggagaattc aggaccttga 2580

gaaatatgtt gaggacacta aaatagatct ctggtcatac aacgcggagc ttcttgttgc 2640

cctggagaac caacatacaa ttgatctaac tgactcagaa atgaacaaac tgtttgaaaa 2700

aacaaagaag caactaaggg aaaatgctga ggatatgggc aatggttgtt tcaaaatata 2760

ccacaaatgt gacaatgcct gcataggatc aatcagaaat ggaacttatg accacgatgt 2820

atacagagat gaagcattaa acaaccggtt ccagatcaag ggagttgagc tgaagtcagg 2880

gtacaaagat tggatcctat ggatttcctt tgccatatca tgtttttgc tttgtgttgc 2940

tttgttgggg ttcatcatgt gggcctgcca aaagggcaac attaggtgca acatttgcat 3000

ttgaaggcct attttcttta gtttgaattt actgttattc ggtgtgcatt tctatgtttg 3060

gtgagcggtt ttctgtgctc agagtgtgtt tattttatgt aatttaattt ctttgtgagc 3120

tcctgtttag caggtcgtcc cttcagcaag gacacaaaaa gattttaatt ttattaaaaa 3180

aaaaaaaaaa aaagaccggg aattcgatat caagcttatc gacctgcaga tcgttcaaac 3240

atttggcaat aaagtttctt aagattgaat cctgttgccg gtcttgcgat gattatcata 3300

taatttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat gacgttattt 3360

atgagatggg tttttatgat tagagtcccg caattataca tttaatacgc gatagaaaac 3420

aaaatatagc gcgcaaacta ggataaatta tcgcgcgcgg tgtcatctat gttactagat 3480

<210> 127
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF**-H2S157.sl-6r

<400> 127
actaaagaaa ataggccttc atatgcagat cctgcactgc agagacccgt 50

<210> 128
<211> 1716
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/H2 Singapore.

<400> 128
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggaccaaat atgcattgga taccatgcca ataattccac agagaaggtc 120

gacacaattc tagagcggaa cgtcactgtg actcatgcca aggacattct tgagaagacc 180

cataacggaa agttatgcaa actaaacgga atccctccac ttgaactagg ggactgtagc 240

attgccggat ggctccttgg aaatccagaa tgtgataggc ttctaagtgt gccagaatgg 300

tcctatataa tggagaaaga aaacccgaga gacggtttgt gttatccagg cagcttcaat 360

gattatgaag aattgaaaca tctcctcagc agcgtgaaac atttcgagaa agtaaagatt 420

ctgcccaaag atagatggac acagcataca acaactggag gttcacgggc ctgcgcggtg 480

tctggtaatc catcattctt caggaacatg gtctggctga caaagaaaga atcaaattat 540

ccggttgcca aaggatcgta caacaataca agcggagaac aaatgctaat aatttggggg 600

gtgcaccatc ccaatgatga gacagaacaa agaacattgt accagaatgt gggaacctat 660

gtttccgtag gcacatcaac attgaacaaa aggtcaaccc cagacatagc aacaaggcct 720

aaagtgaatg gactaggaag tagaatggag ttctcttgga ccctattgga tatgtgggac 780

accataaatt ttgagagtac tggtaatcta attgcaccag agtatggatt caaaatatcg 840

aaaagaggta gttcagggat catgaaaaca gaaggaacac ttgagaactg tgagaccaaa 900

tgccaaactc ctttgggagc aataaataca acattgcctt ttcacaatgt ccacccactg 960

acaataggtg agtgccccaa atatgtaaaa tcggagaagt tggtcttagc aacaggacta 1020

aggaatgttc cccagattga atcaagagga ttgtttgggg caatagctgg ttttatagaa 1080

ggaggatggc aaggaatggt tgatggttgg tatggatacc atcacagcaa tgaccaggga 1140

tcagggtatg cagcagacaa agaatccact caaaaggcat ttgatggaat caccaacaag 1200

gtaaattctg tgattgaaaa gatgaacacc caatttgaag ctgttgggaa agagttcagt 1260

aacttagaga gaagactgga gaacttgaac aaaaagatgg aagacgggtt tctagatgtg 1320

tggacataca atgctgagct tctagttctg atggaaaatg agaggacact tgactttcat 1380

gattctaatg tcaagaatct gtatgataaa gtcagaatgc agctgagaga caacgtcaaa 1440

gaactaggaa atggatgttt tgaattttat cacaaatgtg atgatgaatg catgaatagt 1500

gtgaaaaacg ggacgtatga ttatcccaag tatgaagaag agtctaaact aaatagaaat 1560

gaaatcaaag gggtaaaatt gagcagcatg ggggtttatc aaatccttgc catttatgct 1620

acagtagcag gttctctgtc actggcaatc atgatggctg ggatctcttt ctggatgtgc 1680

tccaacgggt ctctgcagtg caggatctgc atatga 1716

<210> 129
<211> 3471
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cát xét biểu hiện số 2220

<400> 129
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccaatacc gcggagaaaa tggcgaaaaa cgttgcgatt ttcggcttat tgttttctct 1320

tcttgtgttg gttccttctc agatcttcgc ggaccaaata tgcattggat accatgccaa 1380

taattccaca gagaaggtcg acacaattct agagcggaac gtcactgtga ctcatgccaa 1440

ggacattctt gagaagaccc ataacggaaa gttatgcaaa ctaaacggaa tccctccact 1500

tgaactaggg gactgtagca ttgccggatg gctccttgga aatccagaat gtgataggct 1560

tctaagtgtg ccagaatggt cctatataat ggagaaagaa aacccgagag acggtttgtg 1620

ttatccaggc agcttcaatg attatgaaga attgaaacat ctcctcagca gcgtgaaaca 1680

tttcgagaaa gtaaagattc tgcccaaaga tagatggaca cagcatacaa caactggagg 1740

ttcacgggcc tgcgcggtgt ctggtaatcc atcattcttc aggaacatgg tctggctgac 1800

aaagaaagaa tcaaattatc cggttgccaa aggatcgtac aacaatacaa gcggagaaca 1860

aatgctaata atttgggggg tgcaccatcc caatgatgag acagaacaaa gaacattgta 1920

ccagaatgtg ggaacctatg tttccgtagg cacatcaaca ttgaacaaaa ggtcaacccc 1980

agacatagca acaaggccta aagtgaatgg actaggaagt agaatggagt tctcttggac 2040

cctattggat atgtgggaca ccataaattt tgagagtact ggtaatctaa ttgcaccaga 2100

gtatggattc aaaatatcga aaagaggtag ttcagggatc atgaaaacag aaggaacact 2160

tgagaactgt gagaccaaat gccaaactcc tttgggagca ataaatacaa cattgccttt 2220

tcacaatgtc cacccactga caataggtga gtgccccaaa tatgtaaaat cggagaagtt 2280

ggtcttagca acaggactaa ggaatgttcc ccagattgaa tcaagaggat tgtttggggc 2340

aatagctggt tttatagaag gaggatggca aggaatggtt gatggttggt atggatacca 2400

tcacagcaat gaccagggat cagggtatgc agcagacaaa gaatccactc aaaaggcatt 2460

tgatggaatc accaacaagg taaattctgt gattgaaaag atgaacaccc aatttgaagc 2520

tgttgggaaa gagttcagta acttagagag aagactggag aacttgaaca aaaagatgga 2580

agacgggttt ctagatgtgt ggacatacaa tgctgagctt ctagttctga tggaaaatga 2640

gaggacactt gactttcatg attctaatgt caagaatctg tatgataaag tcagaatgca 2700

gctgagagac aacgtcaaag aactaggaaa tggatgtttt gaattttatc acaaatgtga 2760

tgatgaatgc atgaatagtg tgaaaaacgg gacgtatgat tatcccaagt atgaagaaga 2820

gtctaaacta aatagaaatg aaatcaaagg ggtaaaattg agcagcatgg gggtttatca 2880

aatccttgcc atttatgcta cagtagcagg ttctctgtca ctggcaatca tgatggctgg 2940

gatctctttc tggatgtgct ccaacgggtc tctgcagtgc aggatctgca tatgaaggcc 3000

tattttcttt agtttgaatt tactgttatt cggtgtgcat ttctatgttt ggtgagcggt 3060

tttctgtgct cagagtgtgt ttattttatg taatttaatt tctttgtgag ctcctgttta 3120

gcaggtcgtc ccttcagcaa ggacacaaaa agattttaat tttattaaaa aaaaaaaaaa 3180

aaaagaccgg gaattcgata tcaagcttat cgacctgcag atcgttcaaa catttggcaa 3240

taaagtttct taagattgaa tcctgttgcc ggtcttgcga tgattatcat ataatttctg 3300

ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg 3360

gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag 3420

cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga t 3471

<210> 130
<211> 571
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H2 Singapore

<400> 130

```
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                  10                  15

Leu Val Pro Ser Gln Ile Phe Ala Asp Gln Ile Cys Ile Gly Tyr His
            20                  25                  30

Ala Asn Asn Ser Thr Glu Lys Val Asp Thr Ile Leu Glu Arg Asn Val
        35                  40                  45

Thr Val Thr His Ala Lys Asp Ile Leu Glu Lys Thr His Asn Gly Lys
    50                  55                  60

Leu Cys Lys Leu Asn Gly Ile Pro Pro Leu Glu Leu Gly Asp Cys Ser
65                  70                  75                  80

Ile Ala Gly Trp Leu Leu Gly Asn Pro Glu Cys Asp Arg Leu Leu Ser
                85                  90                  95

Val Pro Glu Trp Ser Tyr Ile Met Glu Lys Glu Asn Pro Arg Asp Gly
            100                 105                 110

Leu Cys Tyr Pro Gly Ser Phe Asn Asp Tyr Glu Glu Leu Lys His Leu
        115                 120                 125

Leu Ser Ser Val Lys His Phe Glu Lys Val Lys Ile Leu Pro Lys Asp
    130                 135                 140

Arg Trp Thr Gln His Thr Thr Thr Gly Gly Ser Arg Ala Cys Ala Val
145                 150                 155                 160

Ser Gly Asn Pro Ser Phe Phe Arg Asn Met Val Trp Leu Thr Lys Lys
                165                 170                 175

Glu Ser Asn Tyr Pro Val Ala Lys Gly Ser Tyr Asn Asn Thr Ser Gly
            180                 185                 190

Glu Gln Met Leu Ile Ile Trp Gly Val His His Pro Asn Asp Glu Thr
        195                 200                 205

Glu Gln Arg Thr Leu Tyr Gln Asn Val Gly Thr Tyr Val Ser Val Gly
```

210 215 220

Thr Ser Thr Leu Asn Lys Arg Ser Thr Pro Asp Ile Ala Thr Arg Pro
225 230 235 240

Lys Val Asn Gly Leu Gly Ser Arg Met Glu Phe Ser Trp Thr Leu Leu
245 250 255

Asp Met Trp Asp Thr Ile Asn Phe Glu Ser Thr Gly Asn Leu Ile Ala
260 265 270

Pro Glu Tyr Gly Phe Lys Ile Ser Lys Arg Gly Ser Ser Gly Ile Met
275 280 285

Lys Thr Glu Gly Thr Leu Glu Asn Cys Glu Thr Lys Cys Gln Thr Pro
290 295 300

Leu Gly Ala Ile Asn Thr Thr Leu Pro Phe His Asn Val His Pro Leu
305 310 315 320

Thr Ile Gly Glu Cys Pro Lys Tyr Val Lys Ser Glu Lys Leu Val Leu
325 330 335

Ala Thr Gly Leu Arg Asn Val Pro Gln Ile Glu Ser Arg Gly Leu Phe
340 345 350

Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp
355 360 365

Gly Trp Tyr Gly Tyr His His Ser Asn Asp Gln Gly Ser Gly Tyr Ala
370 375 380

Ala Asp Lys Glu Ser Thr Gln Lys Ala Phe Asp Gly Ile Thr Asn Lys
385 390 395 400

Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Glu Ala Val Gly
405 410 415

Lys Glu Phe Ser Asn Leu Glu Arg Arg Leu Glu Asn Leu Asn Lys Lys
420 425 430

```
Met Glu Asp Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu
        435                 440                 445

Val Leu Met Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val
    450                 455                 460

Lys Asn Leu Tyr Asp Lys Val Arg Met Gln Leu Arg Asp Asn Val Lys
465                 470                 475                 480

Glu Leu Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asp Glu
                485                 490                 495

Cys Met Asn Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Glu
            500                 505                 510

Glu Glu Ser Lys Leu Asn Arg Asn Glu Ile Lys Gly Val Lys Leu Ser
        515                 520                 525

Ser Met Gly Val Tyr Gln Ile Leu Ala Ile Tyr Ala Thr Val Ala Gly
    530                 535                 540

Ser Leu Ser Leu Ala Ile Met Met Ala Gly Ile Ser Phe Trp Met Cys
545                 550                 555                 560

Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
                565                 570
```

<210> 131
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H2S157(Prl-).r

<400> 131
tgccatcctc cgccgggaac attccttagt cctgttgcta agaccaac 48

<210> 132
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H2S157(Prl-).c

<400> 132
aggaatgttc ccggcggagg atggcaagga atggttgatg gttggtatgg 50

<210> 133
<211> 3426
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 2221

<400> 133
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta ataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccaatacc gcggagaaaa tggcgaaaaa cgttgcgatt ttcggcttat tgttttctct 1320

tcttgtgttg gttccttctc agatcttcgc ggaccaaata tgcattggat accatgccaa 1380

taattccaca gagaaggtcg acacaattct agagcggaac gtcactgtga ctcatgccaa 1440

ggacattctt gagaagaccc ataacggaaa gttatgcaaa ctaaacggaa tccctccact 1500

tgaactaggg gactgtagca ttgccggatg gctccttgga aatccagaat gtgataggct 1560

tctaagtgtg ccagaatggt cctatataat ggagaaagaa aacccgagag acggtttgtg 1620

ttatccaggc agcttcaatg attatgaaga attgaaacat ctcctcagca gcgtgaaaca 1680

tttcgagaaa gtaaagattc tgcccaaaga tagatggaca cagcatacaa caactggagg 1740

ttcacgggcc tgcgcggtgt ctggtaatcc atcattcttc aggaacatgg tctggctgac 1800

aaagaaagaa tcaaattatc cggttgccaa aggatcgtac aacaatacaa gcggagaaca 1860

aatgctaata atttgggggg tgcaccatcc caatgatgag acagaacaaa gaacattgta 1920

ccagaatgtg ggaacctatg tttccgtagg cacatcaaca ttgaacaaaa ggtcaacccc 1980

agacatagca acaaggccta aagtgaatgg actaggaagt agaatggagt tctcttggac 2040

cctattggat atgtgggaca ccataaattt tgagagtact ggtaatctaa ttgcaccaga 2100

gtatggattc aaaatatcga aaagaggtag ttcagggatc atgaaaacag aaggaacact 2160

tgagaactgt gagaccaaat gccaaactcc tttgggagca ataaatacaa cattgccttt 2220

tcacaatgtc cacccactga caataggtga gtgccccaaa tatgtaaaat cggagaagtt 2280

ggtcttagca acaggactaa ggaatgttcc cggcggagga tggcaaggaa tggttgatgg 2340

ttggtatgga taccatcaca gcaatgacca gggatcaggg tatgcagcag acaaagaatc 2400

cactcaaaag gcatttgatg gaatcaccaa caaggtaaat tctgtgattg aaaagatgaa 2460

cacccaattt gaagctgttg ggaaagagtt cagtaactta gagagaagac tggagaactt 2520

gaacaaaaag atggaagacg ggtttctaga tgtgtggaca tacaatgctg agcttctagt 2580

tctgatggaa aatgagagga cacttgactt tcatgattct aatgtcaaga atctgtatga 2640

taaagtcaga atgcagctga gagacaacgt caaagaacta ggaaatggat gttttgaatt 2700

ttatcacaaa tgtgatgatg aatgcatgaa tagtgtgaaa aacgggacgt atgattatcc 2760

caagtatgaa gaagagtcta aactaaatag aaatgaaatc aaaggggtaa aattgagcag 2820

catgggggtt tatcaaatcc ttgccattta tgctacagta gcaggttctc tgtcactggc 2880

aatcatgatg gctgggatct ctttctggat gtgctccaac gggtctctgc agtgcaggat 2940

ctgcatatga aggcctattt tctttagttt gaatttactg ttattcggtg tgcatttcta 3000

tgtttggtga gcggttttct gtgctcagag tgtgtttatt ttatgtaatt taatttcttt 3060

gtgagctcct gtttagcagg tcgtcccttc agcaaggaca caaaaagatt ttaattttat 3120

taaaaaaaaa aaaaaaaaag accgggaatt cgatatcaag cttatcgacc tgcagatcgt 3180

tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt 3240

atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg 3300

ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta atacgcgata 3360

gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta 3420

ctagat 3426

<210> 134
<211> 556
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H2 Singapore với vòng phân giải protein loại bớt

<400> 134

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Gln Ile Cys Ile Gly Tyr His
20 25 30

Ala Asn Asn Ser Thr Glu Lys Val Asp Thr Ile Leu Glu Arg Asn Val
35 40 45

Thr Val Thr His Ala Lys Asp Ile Leu Glu Lys Thr His Asn Gly Lys
50 55 60

Leu Cys Lys Leu Asn Gly Ile Pro Pro Leu Glu Leu Gly Asp Cys Ser
65 70 75 80

Ile Ala Gly Trp Leu Leu Gly Asn Pro Glu Cys Asp Arg Leu Leu Ser
85 90 95

Val Pro Glu Trp Ser Tyr Ile Met Glu Lys Glu Asn Pro Arg Asp Gly
100 105 110

Leu Cys Tyr Pro Gly Ser Phe Asn Asp Tyr Glu Glu Leu Lys His Leu
115 120 125

Leu Ser Ser Val Lys His Phe Glu Lys Val Lys Ile Leu Pro Lys Asp
130 135 140

Arg Trp Thr Gln His Thr Thr Thr Gly Gly Ser Arg Ala Cys Ala Val
145 150 155 160

Ser Gly Asn Pro Ser Phe Phe Arg Asn Met Val Trp Leu Thr Lys Lys
165 170 175

Glu Ser Asn Tyr Pro Val Ala Lys Gly Ser Tyr Asn Asn Thr Ser Gly
180 185 190

Glu Gln Met Leu Ile Ile Trp Gly Val His His Pro Asn Asp Glu Thr
195 200 205

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

Thr Ser Thr Leu Asn Lys Arg Ser Thr Pro Asp Ile Ala Thr Arg Pro
225 230 235 240

Lys Val Asn Gly Leu Gly Ser Arg Met Glu Phe Ser Trp Thr Leu Leu
245 250 255

Asp Met Trp Asp Thr Ile Asn Phe Glu Ser Thr Gly Asn Leu Ile Ala
260 265 270

Pro Glu Tyr Gly Phe Lys Ile Ser Lys Arg Gly Ser Ser Gly Ile Met
275 280 285

Lys Thr Glu Gly Thr Leu Glu Asn Cys Glu Thr Lys Cys Gln Thr Pro
290 295 300

Leu Gly Ala Ile Asn Thr Thr Leu Pro Phe His Asn Val His Pro Leu
305 310 315 320

Thr Ile Gly Glu Cys Pro Lys Tyr Val Lys Ser Glu Lys Leu Val Leu
325 330 335

Ala Thr Gly Leu Arg Asn Val Pro Gly Gly Gly Trp Gln Gly Met Val
340 345 350

Asp Gly Trp Tyr Gly Tyr His His Ser Asn Asp Gln Gly Ser Gly Tyr
355 360 365

Ala Ala Asp Lys Glu Ser Thr Gln Lys Ala Phe Asp Gly Ile Thr Asn
370 375 380

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Glu Ala Val
385 390 395 400

Gly Lys Glu Phe Ser Asn Leu Glu Arg Arg Leu Glu Asn Leu Asn Lys
405 410 415

Lys Met Glu Asp Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu
420 425 430

Leu Val Leu Met Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
435 440 445

Val Lys Asn Leu Tyr Asp Lys Val Arg Met Gln Leu Arg Asp Asn Val
450 455 460

Lys Glu Leu Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asp
465 470 475 480

Glu Cys Met Asn Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
485 490 495

Glu Glu Glu Ser Lys Leu Asn Arg Asn Glu Ile Lys Gly Val Lys Leu
500 505 510

Ser Ser Met Gly Val Tyr Gln Ile Leu Ala Ile Tyr Ala Thr Val Ala
515 520 525

Gly Ser Leu Ser Leu Ala Ile Met Met Ala Gly Ile Ser Phe Trp Met
530 535 540

Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
545 550 555

<210> 135
<211> 3120
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 2222

<400> 135
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaggg cccaataccg cggagaaaat ggcgaaaaac gttgcgattt tcggcttatt 960

gttttctctt cttgtgttgg ttccttctca gatcttcgcg gaccaaatat gcattggata 1020

ccatgccaat aattccacag agaaggtcga cacaattcta gagcggaacg tcactgtgac 1080

tcatgccaag gacattcttg agaagaccca taacggaaag ttatgcaaac taaacggaat 1140

ccctccactt gaactagggg actgtagcat tgccggatgg ctccttggaa atccagaatg 1200

tgataggctt ctaagtgtgc cagaatggtc ctatataatg gagaaagaaa acccgagaga 1260

cggtttgtgt tatccaggca gcttcaatga ttatgaagaa ttgaaacatc tcctcagcag 1320

cgtgaaacat ttcgagaaag taaagattct gcccaaagat agatggacac agcatacaac 1380

aactggaggt tcacgggcct gcgcggtgtc tggtaatcca tcattcttca ggaacatggt 1440

ctggctgaca aagaaagaat caaattatcc ggttgccaaa ggatcgtaca acaatacaag 1500

cggagaacaa atgctaataa tttgggggggt gcaccatccc aatgatgaga cagaacaaag 1560

aacattgtac cagaatgtgg gaacctatgt ttccgtaggc acatcaacat tgaacaaaag 1620

gtcaacccca gacatagcaa caaggcctaa agtgaatgga ctaggaagta gaatggagtt 1680

ctcttggacc ctattggata tgtgggacac cataaatttt gagagtactg gtaatctaat 1740

tgcaccagag tatggattca aaatatcgaa aagaggtagt tcagggatca tgaaaacaga 1800

aggaacactt gagaactgtg agaccaaatg ccaaactcct ttgggagcaa taaatacaac 1860

attgcctttt cacaatgtcc acccactgac aataggtgag tgccccaaat atgtaaaatc 1920

ggagaagttg gtcttagcaa caggactaag gaatgttccc cagattgaat caagaggatt 1980

gtttggggca atagctggtt ttatagaagg aggatggcaa ggaatggttg atggttggta 2040

tggataccat cacagcaatg accagggatc agggtatgca gcagacaaag aatccactca 2100

aaaggcattt gatggaatca ccaacaaggt aaattctgtg attgaaaaga tgaacaccca 2160

atttgaagct gttgggaaag agttcagtaa cttagagaga agactggaga acttgaacaa 2220

aaagatggaa gacgggtttc tagatgtgtg gacatacaat gctgagcttc tagttctgat 2280

ggaaaatgag aggacacttg actttcatga ttctaatgtc aagaatctgt atgataaagt 2340

cagaatgcag ctgagagaca acgtcaaaga actaggaaat ggatgttttg aattttatca 2400

caaatgtgat gatgaatgca tgaatagtgt gaaaaacggg acgtatgatt atcccaagta 2460

tgaagaagag tctaaactaa atagaaatga aatcaaaggg gtaaaattga gcagcatggg 2520

ggtttatcaa atccttgcca tttatgctac agtagcaggt tctctgtcac tggcaatcat 2580

gatggctggg atctctttct ggatgtgctc caacgggtct ctgcagtgca ggatctgcat 2640

atgaaggcct attttcttta gtttgaattt actgttattc ggtgtgcatt tctatgtttg 2700

gtgagcggtt ttctgtgctc agagtgtgtt tattttatgt aatttaattt ctttgtgagc 2760

tcctgtttag caggtcgtcc cttcagcaag gacacaaaaa gattttaatt ttattaaaaa 2820

aaaaaaaaaa aaagaccggg aattcgatat caagcttatc gacctgcaga tcgttcaaac 2880

atttggcaat aaagtttctt aagattgaat cctgttgccg gtcttgcgat gattatcata 2940

taatttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat gacgttattt 3000

atgagatggg tttttatgat tagagtcccg caattataca tttaatacgc gatagaaaac 3060

aaaatatagc gcgcaaacta ggataaatta tcgcgcgcgg tgtcatctat gttactagat 3120

<210> 136
<211> 3075
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 2223

<400> 136
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaggg cccaataccg cggagaaaat ggcgaaaaac gttgcgattt tcggcttatt 960

gttttctctt cttgtgttgg ttccttctca gatcttcgcg gaccaaatat gcattggata 1020

ccatgccaat aattccacag agaaggtcga cacaattcta gagcggaacg tcactgtgac 1080

tcatgccaag gacattcttg agaagaccca taacggaaag ttatgcaaac taaacggaat 1140

ccctccactt gaactagggg actgtagcat tgccggatgg ctccttggaa atccagaatg 1200

tgataggctt ctaagtgtgc cagaatggtc ctatataatg gagaaagaaa acccgagaga 1260

cggtttgtgt tatccaggca gcttcaatga ttatgaagaa ttgaaacatc tcctcagcag 1320

cgtgaaacat ttcgagaaag taaagattct gcccaaagat agatggacac agcatacaac 1380

aactggaggt tcacgggcct gcgcggtgtc tggtaatcca tcattcttca ggaacatggt 1440

ctggctgaca aagaaagaat caaattatcc ggttgccaaa ggatcgtaca acaatacaag 1500

cggagaacaa atgctaataa tttgggggggt gcaccatccc aatgatgaga cagaacaaag 1560

aacattgtac cagaatgtgg gaacctatgt ttccgtaggc acatcaacat tgaacaaaag 1620

gtcaacccca gacatagcaa caaggcctaa agtgaatgga ctaggaagta gaatggagtt 1680

ctcttggacc ctattggata tgtgggacac cataaatttt gagagtactg gtaatctaat 1740

tgcaccagag tatggattca aaatatcgaa aagaggtagt tcagggatca tgaaaacaga 1800

aggaacactt gagaactgtg agaccaaatg ccaaactcct ttgggagcaa taaatacaac 1860

attgcctttt cacaatgtcc acccactgac aataggtgag tgccccaaat atgtaaaatc 1920

ggagaagttg gtcttagcaa caggactaag gaatgttccc ggcggaggat ggcaaggaat 1980

ggttgatggt tggtatggat accatcacag caatgaccag ggatcagggt atgcagcaga 2040

caaagaatcc actcaaaagg catttgatgg aatcaccaac aaggtaaatt ctgtgattga 2100

aaagatgaac acccaatttg aagctgttgg gaaagagttc agtaacttag agagaagact 2160

ggagaacttg aacaaaaaga tggaagacgg gtttctagat gtgtggacat acaatgctga 2220

gcttctagtt ctgatggaaa atgagaggac acttgacttt catgattcta atgtcaagaa 2280

tctgtatgat aaagtcagaa tgcagctgag agacaacgtc aaagaactag gaaatggatg 2340

ttttgaattt tatcacaaat gtgatgatga atgcatgaat agtgtgaaaa acgggacgta 2400

tgattatccc aagtatgaag aagagtctaa actaaataga aatgaaatca aaggggtaaa 2460

attgagcagc atgggggttt atcaaatcct tgccatttat gctacagtag caggttctct 2520

gtcactggca atcatgatgg ctgggatctc tttctggatg tgctccaacg ggtctctgca 2580

gtgcaggatc tgcatatgaa ggcctatttt ctttagtttg aatttactgt tattcggtgt 2640

gcatttctat gtttggtgag cggttttctg tgctcagagt gtgtttattt tatgtaattt 2700

aatttctttg tgagctcctg tttagcaggt cgtcccttca gcaaggacac aaaaagattt 2760

taattttatt aaaaaaaaaa aaaaaaaaga ccgggaattc gatatcaagc ttatcgacct 2820

gcagatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt 2880

gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa 2940

tgcatgacgt tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa 3000

tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca 3060

tctatgttac tagat 3075

<210> 137
<211> 1725
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/H3 Perth

<400> 137
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgcaaaaact tcctggaaat gacaacagca cggcaacgct gtgccttggg 120

caccatgcag taccaaacgg aacgatagtg aaaacaatca cgaatgacca aattgaagtt 180

actaatgcta ctgagctggt tcagagttcc tcaacaggtg aaatatgcga cagtcctcat 240

cagatccttg atggaaaaaa ctgcacacta atagatgctc tattgggaga ccctcagtgt 300

gatggcttcc aaaataagaa atgggacctt tttgttgaac gcagcaaagc ctacagcaac 360

tgttaccctt atgatgtgcc ggattatgcc tcccttaggt cactagttgc ctcatccggc 420

acactggagt ttaacaatga aagcttcaat tggactggag tcactcaaaa cggaacaagc 480

tctgcttgca taaggagatc taaaaacagt ttctttagta gattgaattg gttgacccac 540

ttaaacttca atacccagc attgaacgtg actatgccaa acaatgaaca atttgacaaa 600

ttgtacattt gggggggttca ccacccgggt acggacaaag accaaatctt cctgtatgct 660

caagcatcag gaagaatcac agtctctacc aaaagaagcc aacaaaccgt aagcccgaat 720

atcggatcta gacccagagt aaggaatatc cctagcagaa taagcatcta ttggacaata 780

gtaaaaccgg gagacatact tttgattaac agcacaggga atctaattgc tcctaggggt 840

tacttcaaaa tacgaagtgg gaaaagctca ataatgagat cagatgcacc cattggcaaa 900

tgcaattctg aatgcatcac tccaaatgga agcattccca atgacaaacc attccaaaat 960

gtaaacagga tcacatacgg ggcctgtccc agatatgtta agcaaaacac tctgaaattg 1020

gcaacaggga tgcgaaatgt accagagaaa caaactagag gcatatttgg cgcaatcgcg 1080

ggtttcatag aaaatggttg ggagggaatg gtggatggtt ggtacggttt caggcatcaa 1140

aattctgagg gaagaggaca agcagcagat ctcaaaagca ctcaagcagc aatcgatcaa 1200

atcaatggga agctgaatag attgatcggg aaaaccaacg agaaattcca tcagattgaa 1260

aaagaattct cagaagtcga agggagaatt caggaccttg agaaatatgt tgaggacact 1320

aaaatagatc tctggtcata caacgcggag cttcttgttg ccctggagaa ccaacataca 1380

attgatctaa ctgactcaga aatgaacaaa ctgtttgaaa aaacaaagaa gcaactgagg 1440

gaaaatgctg aggatatggg caatggttgt ttcaaaatat accacaaatg tgacaatgcc 1500

tgcataggat caatcagaaa tggaacttat gaccacgatg tatacagaga tgaagcatta 1560

aacaaccggt ttcagatcaa gggagttgag ctgaagtcag ggtacaaaga ttggatccta 1620

tggatttcct ttgccatatc atgtttttg ctttgtgttg ctttgttggg gttcatcatg 1680

tgggcctgcc aaaaaggcaa cattaggtgc aacatttgca tttga 1725

<210> 138
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF**-H3P1609.S1-6r

<400> 138
actaaagaaa ataggccttc aaatgcaaat gttgcaccta atgttgcctt 50

<210> 139
<211> 574

<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H3 Perth

<400> 139

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Lys Leu Pro Gly Asn Asp Asn
20 25 30

Ser Thr Ala Thr Leu Cys Leu Gly His His Ala Val Pro Asn Gly Thr
35 40 45

Ile Val Lys Thr Ile Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr
50 55 60

Glu Leu Val Gln Ser Ser Ser Thr Gly Glu Ile Cys Asp Ser Pro His
65 70 75 80

Gln Ile Leu Asp Gly Lys Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly
85 90 95

Asp Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val
100 105 110

Glu Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp
115 120 125

Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe
130 135 140

Asn Asn Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser
145 150 155 160

Ser Ala Cys Ile Arg Arg Ser Lys Asn Ser Phe Phe Ser Arg Leu Asn
165 170 175

Trp Leu Thr His Leu Asn Phe Lys Tyr Pro Ala Leu Asn Val Thr Met
180 185 190

```
Pro Asn Asn Glu Gln Phe Asp Lys Leu Tyr Ile Trp Gly Val His His
        195                 200                 205

Pro Gly Thr Asp Lys Asp Gln Ile Phe Leu Tyr Ala Gln Ala Ser Gly
    210                 215                 220

Arg Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Thr Val Ser Pro Asn
225                 230                 235                 240

Ile Gly Ser Arg Pro Arg Val Arg Asn Ile Pro Ser Arg Ile Ser Ile
                245                 250                 255

Tyr Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr
            260                 265                 270

Gly Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys
        275                 280                 285

Ser Ser Ile Met Arg Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu
    290                 295                 300

Cys Ile Thr Pro Asn Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn
305                 310                 315                 320

Val Asn Arg Ile Thr Tyr Gly Ala Cys Pro Arg Tyr Val Lys Gln Asn
                325                 330                 335

Thr Leu Lys Leu Ala Thr Gly Met Arg Asn Val Pro Glu Lys Gln Thr
            340                 345                 350

Arg Gly Ile Phe Gly Ala Ile Ala Gly Phe Ile Glu Asn Gly Trp Glu
        355                 360                 365

Gly Met Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu Gly
    370                 375                 380

Arg Gly Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp Gln
385                 390                 395                 400
```

Ile Asn Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys Phe
405 410 415

His Gln Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln Asp
420 425 430

Leu Glu Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr Asn
435 440 445

Ala Glu Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu Thr
450 455 460

Asp Ser Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu Arg
465 470 475 480

Glu Asn Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His Lys
485 490 495

Cys Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp His
500 505 510

Asp Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly
515 520 525

Val Glu Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp Ile Ser Phe
530 535 540

Ala Ile Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly Phe Ile Met
545 550 555 560

Trp Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
565 570

<210> 140
<211> 1680
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/H3 Perth với vòng phân giải protein loại bớt

<400> 140

atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgcaaaaact tcctggaaat gacaacagca cggcaacgct gtgccttggg 120

caccatgcag taccaaacgg aacgatagtg aaaacaatca cgaatgacca aattgaagtt 180

actaatgcta ctgagctggt tcagagttcc tcaacaggtg aaatatgcga cagtcctcat 240

cagatccttg atggaaaaaa ctgcacacta atagatgctc tattgggaga ccctcagtgt 300

gatggcttcc aaaataagaa atgggacctt tttgttgaac gcagcaaagc ctacagcaac 360

tgttaccctt atgatgtgcc ggattatgcc tcccttaggt cactagttgc ctcatccggc 420

acactggagt ttaacaatga aagcttcaat tggactggag tcactcaaaa cggaacaagc 480

tctgcttgca taaggagatc taaaaacagt ttctttagta gattgaattg gttgacccac 540

ttaaacttca aatacccagc attgaacgtg actatgccaa acaatgaaca atttgacaaa 600

ttgtacattt ggggggttca ccacccgggt acggacaaag accaaatctt cctgtatgct 660

caagcatcag gaagaatcac agtctctacc aaaagaagcc aacaaaccgt aagcccgaat 720

atcggatcta gacccagagt aaggaatatc cctagcagaa taagcatcta ttggacaata 780

gtaaaaccgg gagacatact tttgattaac agcacaggga atctaattgc tcctaggggt 840

tacttcaaaa tacgaagtgg gaaaagctca ataatgagat cagatgcacc cattggcaaa 900

tgcaattctg aatgcatcac tccaaatgga agcattccca atgacaaacc attccaaaat 960

gtaaacagga tcacatacgg ggcctgtccc agatatgtta agcaaaacac tctgaaattg 1020

gcaacaggga tgcgaaatgt accaggcgga ggttgggagg gaatggtgga tggttggtac 1080

ggtttcaggc atcaaaattc tgagggaaga ggacaagcag cagatctcaa aagcactcaa 1140

gcagcaatcg atcaaatcaa tgggaagctg aatagattga tcgggaaaac caacgagaaa 1200

ttccatcaga ttgaaaaaga attctcagaa gtcgaaggga gaattcagga ccttgagaaa 1260

tatgttgagg acactaaaat agatctctgg tcatacaacg cggagcttct tgttgccctg 1320

gagaaccaac atacaattga tctaactgac tcagaaatga acaaactgtt tgaaaaaaca 1380

aagaagcaac tgagggaaaa tgctgaggat atgggcaatg gttgtttcaa aatataccac 1440

aaatgtgaca atgcctgcat aggatcaatc agaaatggaa cttatgacca cgatgtatac 1500

agagatgaag cattaaacaa ccggtttcag atcaagggag ttgagctgaa gtcagggtac 1560

aaagattgga tcctatggat ttcctttgcc atatcatgtt ttttgctttg tgttgctttg 1620

ttggggttca tcatgtgggc ctgccaaaaa ggcaacatta ggtgcaacat ttgcatttga 1680

<210> 141
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H3P1609(Prl-)#2.r

<400> 141
ccctcccaac ctccgcctgg tacatttcgc atccctgttg ccaatttc 48

<210> 142
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H3P1609(Prl-)#2.c

<400> 142
aatgtaccag gcggaggttg ggagggaatg gtggatggtt ggtacggt 48

<210> 143
<211> 559
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H3 Perth với vòng phân giải protein loại bớt

<400> 143

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Lys Leu Pro Gly Asn Asp Asn
20 25 30

Ser Thr Ala Thr Leu Cys Leu Gly His His Ala Val Pro Asn Gly Thr
35 40 45

Ile Val Lys Thr Ile Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr
50 55 60

```
Glu Leu Val Gln Ser Ser Ser Thr Gly Glu Ile Cys Asp Ser Pro His
65                  70                  75                  80

Gln Ile Leu Asp Gly Lys Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly
                85                  90                  95

Asp Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val
            100                 105                 110

Glu Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp
        115                 120                 125

Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe
    130                 135                 140

Asn Asn Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser
145                 150                 155                 160

Ser Ala Cys Ile Arg Arg Ser Lys Asn Ser Phe Phe Ser Arg Leu Asn
                165                 170                 175

Trp Leu Thr His Leu Asn Phe Lys Tyr Pro Ala Leu Asn Val Thr Met
            180                 185                 190

Pro Asn Asn Glu Gln Phe Asp Lys Leu Tyr Ile Trp Gly Val His His
        195                 200                 205

Pro Gly Thr Asp Lys Asp Gln Ile Phe Leu Tyr Ala Gln Ala Ser Gly
    210                 215                 220

Arg Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Thr Val Ser Pro Asn
225                 230                 235                 240

Ile Gly Ser Arg Pro Arg Val Arg Asn Ile Pro Ser Arg Ile Ser Ile
                245                 250                 255

Tyr Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr
            260                 265                 270

Gly Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys
        275                 280                 285
```

```
Ser Ser Ile Met Arg Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu
    290                 295                 300

Cys Ile Thr Pro Asn Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn
305                 310                 315                 320

Val Asn Arg Ile Thr Tyr Gly Ala Cys Pro Arg Tyr Val Lys Gln Asn
                325                 330                 335

Thr Leu Lys Leu Ala Thr Gly Met Arg Asn Val Pro Gly Gly Gly Trp
            340                 345                 350

Glu Gly Met Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu
        355                 360                 365

Gly Arg Gly Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp
    370                 375                 380

Gln Ile Asn Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys
385                 390                 395                 400

Phe His Gln Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln
                405                 410                 415

Asp Leu Glu Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr
            420                 425                 430

Asn Ala Glu Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu
        435                 440                 445

Thr Asp Ser Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu
    450                 455                 460

Arg Glu Asn Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His
465                 470                 475                 480

Lys Cys Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp
                485                 490                 495
```

His Asp Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys
500 505 510

Gly Val Glu Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp Ile Ser
515 520 525

Phe Ala Ile Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly Phe Ile
530 535 540

Met Trp Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
545 550 555

<210> 144
<211> 1680
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/H3 Victoria với vòng phân giải protein loại bớt

<400> 144
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cgcaaaaact tcctggaaat gacaacagca cggcaacgct gtgccttggg 120

caccatgcag taccaaacgg aacgatagtg aaaacaatca cgaatgacca aattgaagtt 180

actaatgcta ctgagctggt tcagaattcc tcaataggtg aaatatgcga cagtcctcat 240

cagatccttg atggagaaaa ctgcacacta atagatgctc tattgggaga ccctcagtgt 300

gatggcttcc aaaataagaa atgggacctt tttgttgaac gaagcaaagc ctacagcaac 360

tgttaccctt atgatgtgcc ggattatgcc tcccttaggt cactagttgc ctcatccggc 420

acactggagt ttaacaatga aagcttcaat tggactggag tcactcaaaa cggaacaagt 480

tctgcttgca taaggagatc taataatagt ttctttagta gattaaattg gttgacccac 540

ttaaacttca aatacccagc attgaacgtg actatgccaa acaatgaaca atttgacaaa 600

ttgtacattt ggggggttca ccacccgggt acggacaagg accaaatctt cctgtatgct 660

caatcatcag gaagaatcac agtatctacc aaaagaagcc aacaagctgt aatcccgaat 720

atcggatcta gacccagaat aaggaatatc cctagcagaa taagcatcta ttggacaata 780

gtaaaaccgg gagacatact tttgattaac agcacaggga atctaattgc tcctaggggt 840

tacttcaaaa tacgaagtgg gaaaagctca ataatgagat cagatgcacc cattggcaaa 900

tgcaattctg aatgcatcac tccaaatgga agcattccca atgacaaacc attccaaaat 960

gtaaacagga tcacatacgg ggcctgtccc agatatgtta agcaaagcac tctgaaattg 1020

gcaacaggaa tgcgaaatgt accaggcgga ggttgggagg gaatggtgga tggttggtac 1080

ggtttcaggc atcaaaattc tgagggaaga ggacaagcag cagatctcaa aagcactcaa 1140

gcagcaatcg atcaaatcaa tgggaagctg aatcgattga tcgggaaaac caacgagaaa 1200

ttccatcaga ttgaaaaaga attctcagaa gtcgaaggga gaattcagga ccttgagaaa 1260

tatgttgagg acactaaaat agatctctgg tcatacaacg cggagcttct tgttgccctg 1320

gagaaccaac atacaattga tctaactgac tcagaaatga acaaactgtt tgaaaaaaca 1380

aagaagcaac taagggaaaa tgctgaggat atgggcaatg gttgtttcaa aatataccac 1440

aaatgtgaca atgcctgcat aggatcaatc agaaatggaa cttatgacca cgatgtatac 1500

agagatgaag cattaaacaa ccggttccag atcaagggag ttgagctgaa gtcagggtac 1560

aaagattgga tcctatggat ttcctttgcc atatcatgtt ttttgctttg tgttgctttg 1620

ttggggttca tcatgtgggc ctgccaaaag ggcaacatta ggtgcaacat ttgcatttga 1680

<210> 145
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H3V36111(Prl-).r

<400> 145
ccctcccaac ctccgcctgg tacatttcgc attcctgttg ccaatttc 48

<210> 146
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H3V36111(Prl-).c

<400> 146
aatgtaccag gcggaggttg ggagggaatg gtggatggtt ggtacggt 48

<210> 147

<211> 559
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H3 Victoria với vòng phân giải protein loại bớt

<400> 147

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Gln Lys Leu Pro Gly Asn Asp Asn
20 25 30

Ser Thr Ala Thr Leu Cys Leu Gly His His Ala Val Pro Asn Gly Thr
35 40 45

Ile Val Lys Thr Ile Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr
50 55 60

Glu Leu Val Gln Asn Ser Ser Ile Gly Glu Ile Cys Asp Ser Pro His
65 70 75 80

Gln Ile Leu Asp Gly Glu Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly
85 90 95

Asp Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val
100 105 110

Glu Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp
115 120 125

Tyr Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe
130 135 140

Asn Asn Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser
145 150 155 160

Ser Ala Cys Ile Arg Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn
165 170 175

```
Trp Leu Thr His Leu Asn Phe Lys Tyr Pro Ala Leu Asn Val Thr Met
            180                 185                 190

Pro Asn Asn Glu Gln Phe Asp Lys Leu Tyr Ile Trp Gly Val His His
        195                 200                 205

Pro Gly Thr Asp Lys Asp Gln Ile Phe Leu Tyr Ala Gln Ser Ser Gly
    210                 215                 220

Arg Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Ala Val Ile Pro Asn
225                 230                 235                 240

Ile Gly Ser Arg Pro Arg Ile Arg Asn Ile Pro Ser Arg Ile Ser Ile
                245                 250                 255

Tyr Trp Thr Ile Val Lys Pro Gly Asp Ile Leu Leu Ile Asn Ser Thr
            260                 265                 270

Gly Asn Leu Ile Ala Pro Arg Gly Tyr Phe Lys Ile Arg Ser Gly Lys
        275                 280                 285

Ser Ser Ile Met Arg Ser Asp Ala Pro Ile Gly Lys Cys Asn Ser Glu
    290                 295                 300

Cys Ile Thr Pro Asn Gly Ser Ile Pro Asn Asp Lys Pro Phe Gln Asn
305                 310                 315                 320

Val Asn Arg Ile Thr Tyr Gly Ala Cys Pro Arg Tyr Val Lys Gln Ser
                325                 330                 335

Thr Leu Lys Leu Ala Thr Gly Met Arg Asn Val Pro Gly Gly Gly Trp
            340                 345                 350

Glu Gly Met Val Asp Gly Trp Tyr Gly Phe Arg His Gln Asn Ser Glu
        355                 360                 365

Gly Arg Gly Gln Ala Ala Asp Leu Lys Ser Thr Gln Ala Ala Ile Asp
    370                 375                 380

Gln Ile Asn Gly Lys Leu Asn Arg Leu Ile Gly Lys Thr Asn Glu Lys
385                 390                 395                 400
```

Phe His Gln Ile Glu Lys Glu Phe Ser Glu Val Glu Gly Arg Ile Gln
405 410 415

Asp Leu Glu Lys Tyr Val Glu Asp Thr Lys Ile Asp Leu Trp Ser Tyr
420 425 430

Asn Ala Glu Leu Leu Val Ala Leu Glu Asn Gln His Thr Ile Asp Leu
435 440 445

Thr Asp Ser Glu Met Asn Lys Leu Phe Glu Lys Thr Lys Lys Gln Leu
450 455 460

Arg Glu Asn Ala Glu Asp Met Gly Asn Gly Cys Phe Lys Ile Tyr His
465 470 475 480

Lys Cys Asp Asn Ala Cys Ile Gly Ser Ile Arg Asn Gly Thr Tyr Asp
485 490 495

His Asp Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys
500 505 510

Gly Val Glu Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp Ile Ser
515 520 525

Phe Ala Ile Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly Phe Ile
530 535 540

Met Trp Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
545 550 555

<210> 148
<211> 1701
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/H7 Hangzhou

<400> 148
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggacaaaat ctgcctcgga catcatgccg tgtcaaacgg aaccaaagta 120

aacacattaa ctgaaagagg agtggaagtc gtcaatgcaa ctgaaacagt ggaacgaaca 180

aacatcccca ggatctgctc aaaagggaaa aggacagttg acctcggtca atgtggactc 240

ctggggacaa tcactggacc acctcaatgt gaccaattcc tagaattttc agccgattta 300

attattgaga ggcgagaagg aagtgatgtc tgttatcctg ggaaattcgt gaatgaagaa 360

gctctgaggc aaattctcag agaatcaggc ggaattgaca aggaagcaat gggattcaca 420

tacagtggaa taagaactaa tggagcaacc agtgcatgta ggagatcagg atcttcattc 480

tatgcagaaa tgaaatggct cctgtcaaac acagataatg ctgcattccc gcagatgact 540

aagtcatata aaaatacaag aaaaagccca gctctaatag tatgggggat ccatcattcc 600

gtatcaactg cagagcaaac caagctatat gggagtggaa acaaactggt gacagttggg 660

agttctaatt atcaacaatc ttttgtaccg agtccaggag cgagaccaca agttaatggt 720

atatctggaa gaattgactt tcattggcta atgctaaatc ccaatgatac agtcactttc 780

agtttcaatg gggctttcat agctccagac cgtgcaagct tcctgagagg aaaatctatg 840

ggaatccaga gtggagtaca ggttgatgcc aattgtgaag gggactgcta tcatagtgga 900

gggacaataa taagtaactt gccatttcag aacatagata gcagggcagt tggaaaatgt 960

ccgagatatg ttaagcaaag gagtctgctg ctagcaacag ggatgaagaa tgttcctgag 1020

attccaaagg gaagaggcct atttggtgct atagcgggtt tcattgaaaa tggatgggaa 1080

ggcctaattg atggttggta tggtttcaga caccagaatg cacagggaga gggaactgct 1140

gcagattaca aaagcactca atcggcaatt gatcaaataa caggaaaatt aaaccggctt 1200

atagaaaaaa ccaaccaaca atttgagttg atcgacaatg aattcaatga ggtagagaag 1260

caaatcggta atgtgataaa ttggaccaga gattctataa cagaagtgtg gtcatacaat 1320

gctgaactct tggtagcaat ggagaaccag catacaattg atctggctga ttcagaaatg 1380

gacaaactgt acgaacgagt gaaaagacag ctgagagaga atgctgaaga agatggcact 1440

ggttgctttg aaatatttca caagtgtgat gatgactgta tggccagtat tagaaataac 1500

acctatgatc acagcaaata cagggaagag gcaatgcaaa atagaataca gattgaccca 1560

gtcaaactaa gcagcggcta caaagatgtg atactttggt ttagcttcgg ggcatcatgt 1620

ttcatacttc tagccattgt aatgggcctt gtcttcatat gtgtaaagaa tggaaacatg 1680

cggtgcacta tttgtatata a 1701

<210> 149
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF*-H7H113.s1-6r

<400> 149
actaaagaaa ataggccttt atatacaaat agtgcaccgc atgtttccat 50

<210> 150
<211> 566
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H7 Hangzhou

<400> 150

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Lys Ile Cys Leu Gly His His
20 25 30

Ala Val Ser Asn Gly Thr Lys Val Asn Thr Leu Thr Glu Arg Gly Val
35 40 45

Glu Val Val Asn Ala Thr Glu Thr Val Glu Arg Thr Asn Ile Pro Arg
50 55 60

Ile Cys Ser Lys Gly Lys Arg Thr Val Asp Leu Gly Gln Cys Gly Leu
65 70 75 80

Leu Gly Thr Ile Thr Gly Pro Pro Gln Cys Asp Gln Phe Leu Glu Phe
85 90 95

Ser Ala Asp Leu Ile Ile Glu Arg Arg Glu Gly Ser Asp Val Cys Tyr
100 105 110

Pro Gly Lys Phe Val Asn Glu Glu Ala Leu Arg Gln Ile Leu Arg Glu
115 120 125

Ser Gly Gly Ile Asp Lys Glu Ala Met Gly Phe Thr Tyr Ser Gly Ile
130 135 140

Arg Thr Asn Gly Ala Thr Ser Ala Cys Arg Arg Ser Gly Ser Ser Phe
145 150 155 160

Tyr Ala Glu Met Lys Trp Leu Leu Ser Asn Thr Asp Asn Ala Ala Phe
165 170 175

Pro Gln Met Thr Lys Ser Tyr Lys Asn Thr Arg Lys Ser Pro Ala Leu
180 185 190

Ile Val Trp Gly Ile His His Ser Val Ser Thr Ala Glu Gln Thr Lys
195 200 205

Leu Tyr Gly Ser Gly Asn Lys Leu Val Thr Val Gly Ser Ser Asn Tyr
210 215 220

Gln Gln Ser Phe Val Pro Ser Pro Gly Ala Arg Pro Gln Val Asn Gly
225 230 235 240

Ile Ser Gly Arg Ile Asp Phe His Trp Leu Met Leu Asn Pro Asn Asp
245 250 255

Thr Val Thr Phe Ser Phe Asn Gly Ala Phe Ile Ala Pro Asp Arg Ala
260 265 270

Ser Phe Leu Arg Gly Lys Ser Met Gly Ile Gln Ser Gly Val Gln Val
275 280 285

Asp Ala Asn Cys Glu Gly Asp Cys Tyr His Ser Gly Gly Thr Ile Ile
290 295 300

Ser Asn Leu Pro Phe Gln Asn Ile Asp Ser Arg Ala Val Gly Lys Cys
305 310 315 320

Pro Arg Tyr Val Lys Gln Arg Ser Leu Leu Leu Ala Thr Gly Met Lys
325 330 335

Asn Val Pro Glu Ile Pro Lys Gly Arg Gly Leu Phe Gly Ala Ile Ala
340 345 350

Gly Phe Ile Glu Asn Gly Trp Glu Gly Leu Ile Asp Gly Trp Tyr Gly
355 360 365

Phe Arg His Gln Asn Ala Gln Gly Glu Gly Thr Ala Ala Asp Tyr Lys
370 375 380

Ser Thr Gln Ser Ala Ile Asp Gln Ile Thr Gly Lys Leu Asn Arg Leu
385 390 395 400

Ile Glu Lys Thr Asn Gln Gln Phe Glu Leu Ile Asp Asn Glu Phe Asn
405 410 415

Glu Val Glu Lys Gln Ile Gly Asn Val Ile Asn Trp Thr Arg Asp Ser
420 425 430

Ile Thr Glu Val Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Met Glu
435 440 445

Asn Gln His Thr Ile Asp Leu Ala Asp Ser Glu Met Asp Lys Leu Tyr
450 455 460

Glu Arg Val Lys Arg Gln Leu Arg Glu Asn Ala Glu Glu Asp Gly Thr
465 470 475 480

Gly Cys Phe Glu Ile Phe His Lys Cys Asp Asp Asp Cys Met Ala Ser
485 490 495

Ile Arg Asn Asn Thr Tyr Asp His Ser Lys Tyr Arg Glu Glu Ala Met
500 505 510

Gln Asn Arg Ile Gln Ile Asp Pro Val Lys Leu Ser Ser Gly Tyr Lys
515 520 525

Asp Val Ile Leu Trp Phe Ser Phe Gly Ala Ser Cys Phe Ile Leu Leu
530 535 540

Ala Ile Val Met Gly Leu Val Phe Ile Cys Val Lys Asn Gly Asn Met
545 550 555 560

Arg Cys Thr Ile Cys Ile
565

<210> 151
<211> 1650
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/H7 Hangzhou với vòng phân giải protein loại bớt

<400> 151
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggacaaaat ctgcctcgga catcatgccg tgtcaaacgg aaccaaagta 120

aacacattaa ctgaaagagg agtggaagtc gtcaatgcaa ctgaaacagt ggaacgaaca 180

aacatcccca ggatctgctc aaaagggaaa aggacagttg acctcggtca atgtggactc 240

ctggggacaa tcactggacc acctcaatgt gaccaattcc tagaattttc agccgattta 300

attattgaga ggcgagaagg aagtgatgtc tgttatcctg ggaaattcgt gaatgaagaa 360

gctctgaggc aaattctcag agaatcaggc ggaattgaca aggaagcaat gggattcaca 420

tacagtggaa taagaactaa tggagcaacc agtgcatgta ggagatcagg atcttcattc 480

tatgcagaaa tgaaatggct cctgtcaaac acagataatg ctgcattccc gcagatgact 540

aagtcatata aaaatacaag aaaaagccca gctctaatag tatgggggat ccatcattcc 600

gtatcaactg cagagcaaac caagctatat gggagtggaa acaaactggt gacagttggg 660

agttctaatt atcaacaatc ttttgtaccg agtccaggag cgagaccaca agttaatggt 720

atatctggaa gaattgactt tcattggcta atgctaaatc ccaatgatac agtcactttc 780

agtttcaatg gggctttcat agctccagac cgtgcaagct tcctgagagg aaaatctatg 840

ggaatccaga gtggagtaca ggttgatgcc aattgtgaag gggactgcta tcatagtgga 900

gggacaataa taagtaactt gccatttcag aacatagata gcagggcagt tggaaaatgt 960

ccgagatatg ttaagcaaag gagtctgctg ctagcaacag ggatgaagaa tgttcctggc 1020

ggatgggaag gcctaattga tggttggtat ggtttcagac accagaatgc acagggagag 1080

ggaactgctg cagattacaa aagcactcaa tcggcaattg atcaaataac aggaaaatta 1140

aaccggctta tagaaaaaac caaccaacaa tttgagttga tcgacaatga attcaatgag 1200

gtagagaagc aaatcggtaa tgtgataaat tggaccagag attctataac agaagtgtgg 1260

tcatacaatg ctgaactctt ggtagcaatg gagaaccagc atacaattga tctggctgat 1320

tcagaaatgg acaaactgta cgaacgagtg aaaagacagc tgagagagaa tgctgaagaa 1380

gatggcactg gttgctttga aatatttcac aagtgtgatg atgactgtat ggccagtatt 1440

agaaataaca cctatgatca cagcaaatac agggaagagg caatgcaaaa tagaatacag 1500

attgacccag tcaaactaag cagcggctac aaagatgtga tactttggtt tagcttcggg 1560

gcatcatgtt tcatacttct agccattgta atgggccttg tcttcatatg tgtaaagaat 1620

ggaaacatgc ggtgcactat ttgtatataa 1650

<210> 152
<211> 49
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H7H113(PrL-).r

<400> 152
ccttcccatc cgccaggaac attcttcatc cctgttgcta gcagcagac 49

<210> 153
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H7H113(PrL-).c

<400> 153
agaatgttcc tggcggatgg gaaggcctaa ttgatggttg gtatggtt 48

<210> 154
<211> 549
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H7 Hangzhou với vòng phân giải protein loại bớt

<400> 154

```
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                  10                  15

Leu Val Pro Ser Gln Ile Phe Ala Asp Lys Ile Cys Leu Gly His His
            20                  25                  30

Ala Val Ser Asn Gly Thr Lys Val Asn Thr Leu Thr Glu Arg Gly Val
        35                  40                  45

Glu Val Val Asn Ala Thr Glu Thr Val Glu Arg Thr Asn Ile Pro Arg
    50                  55                  60

Ile Cys Ser Lys Gly Lys Arg Thr Val Asp Leu Gly Gln Cys Gly Leu
65                  70                  75                  80

Leu Gly Thr Ile Thr Gly Pro Pro Gln Cys Asp Gln Phe Leu Glu Phe
                85                  90                  95

Ser Ala Asp Leu Ile Ile Glu Arg Arg Glu Gly Ser Asp Val Cys Tyr
            100                 105                 110

Pro Gly Lys Phe Val Asn Glu Glu Ala Leu Arg Gln Ile Leu Arg Glu
        115                 120                 125

Ser Gly Gly Ile Asp Lys Glu Ala Met Gly Phe Thr Tyr Ser Gly Ile
    130                 135                 140

Arg Thr Asn Gly Ala Thr Ser Ala Cys Arg Arg Ser Gly Ser Ser Phe
145                 150                 155                 160

Tyr Ala Glu Met Lys Trp Leu Leu Ser Asn Thr Asp Asn Ala Ala Phe
                165                 170                 175

Pro Gln Met Thr Lys Ser Tyr Lys Asn Thr Arg Lys Ser Pro Ala Leu
            180                 185                 190

Ile Val Trp Gly Ile His His Ser Val Ser Thr Ala Glu Gln Thr Lys
        195                 200                 205

Leu Tyr Gly Ser Gly Asn Lys Leu Val Thr Val Gly Ser Ser Asn Tyr
    210                 215                 220
```

```
Gln Gln Ser Phe Val Pro Ser Pro Gly Ala Arg Pro Gln Val Asn Gly
225                 230                 235                 240

Ile Ser Gly Arg Ile Asp Phe His Trp Leu Met Leu Asn Pro Asn Asp
                245                 250                 255

Thr Val Thr Phe Ser Phe Asn Gly Ala Phe Ile Ala Pro Asp Arg Ala
            260                 265                 270

Ser Phe Leu Arg Gly Lys Ser Met Gly Ile Gln Ser Gly Val Gln Val
        275                 280                 285

Asp Ala Asn Cys Glu Gly Asp Cys Tyr His Ser Gly Gly Thr Ile Ile
    290                 295                 300

Ser Asn Leu Pro Phe Gln Asn Ile Asp Ser Arg Ala Val Gly Lys Cys
305                 310                 315                 320

Pro Arg Tyr Val Lys Gln Arg Ser Leu Leu Leu Ala Thr Gly Met Lys
                325                 330                 335

Asn Val Pro Gly Gly Trp Glu Gly Leu Ile Asp Gly Trp Tyr Gly Phe
            340                 345                 350

Arg His Gln Asn Ala Gln Gly Glu Gly Thr Ala Ala Asp Tyr Lys Ser
        355                 360                 365

Thr Gln Ser Ala Ile Asp Gln Ile Thr Gly Lys Leu Asn Arg Leu Ile
    370                 375                 380

Glu Lys Thr Asn Gln Gln Phe Glu Leu Ile Asp Asn Glu Phe Asn Glu
385                 390                 395                 400

Val Glu Lys Gln Ile Gly Asn Val Ile Asn Trp Thr Arg Asp Ser Ile
                405                 410                 415

Thr Glu Val Trp Ser Tyr Asn Ala Glu Leu Leu Val Ala Met Glu Asn
            420                 425                 430
```

Gln His Thr Ile Asp Leu Ala Asp Ser Glu Met Asp Lys Leu Tyr Glu
435 440 445

Arg Val Lys Arg Gln Leu Arg Glu Asn Ala Glu Glu Asp Gly Thr Gly
450 455 460

Cys Phe Glu Ile Phe His Lys Cys Asp Asp Asp Cys Met Ala Ser Ile
465 470 475 480

Arg Asn Asn Thr Tyr Asp His Ser Lys Tyr Arg Glu Glu Ala Met Gln
485 490 495

Asn Arg Ile Gln Ile Asp Pro Val Lys Leu Ser Ser Gly Tyr Lys Asp
500 505 510

Val Ile Leu Trp Phe Ser Phe Gly Ala Ser Cys Phe Ile Leu Leu Ala
515 520 525

Ile Val Met Gly Leu Val Phe Ile Cys Val Lys Asn Gly Asn Met Arg
530 535 540

Cys Thr Ile Cys Ile
545

<210> 155
<211> 1701
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/H9 Hong Kong

<400> 155
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggataaaat ctgcatcggc caccagtcaa caaactccac agaaactgtg 120

gacacgctaa cagaaaccaa tgttcctgtg acacatgcca aagaattgct ccacacagag 180

cataatggaa tgctgtgtgc aacaagcctg ggacatcccc tcattctaga cacatgcact 240

attgaaggac tagtctatgg caacccttct tgtgacctgc tgttgggagg aagagaatgg 300

tcctacatcg tcgaaagatc atcagctgta aatggaacgt gttaccctgg gaatgtagaa 360

aacctagagg aactcaggac acttttttagt tccgctagtt cctaccaaag aatccaaatc 420

ttcccagaca caacctggaa tgtgacttac actggaacaa gcagagcatg ttcaggttca 480

ttctacagga gtatgagatg gctgactcaa aagagcggtt tttaccctgt tcaagacgcc 540

caatacacaa ataacagggg aaagagcatt cttttcgtgt ggggcataca tcacccaccc 600

acctataccg agcaaacaaa tttgtacata agaaacgaca caacaacaag cgtgacaaca 660

gaagatttga ataggacctt caaaccagtg atagggccaa ggccccttgt caatggtctg 720

cagggaagaa ttgattatta ttggtcggta ctaaaaccag gccaaacatt gcgagtacga 780

tccaatggga atctaattgc tccatggtat ggacacgttc tttcaggagg gagccatgga 840

agaatcctga agactgattt aaaaggtggt aattgtgtag tgcaatgtca gactgaaaaa 900

ggtggcttaa acagtacatt gccattccac aatatcagta aatatgcatt tggaacctgc 960

cccaaatatg taagagttaa tagtctcaaa ctggcagtcg gtctgaggaa cgtgcctgct 1020

agatcaagta gaggactatt tggagccata gctggattca tagaaggagg ttggccagga 1080

ctagtcgctg gctggtatgg tttccagcat tcaaatgatc aaggggttgg tatggctgca 1140

gatagggatt caactcaaaa ggcaattgat aaaataacat ccaaggtgaa taatatagtc 1200

gacaagatga acaagcaata tgaaataatt gatcatgaat ttagtgaggt tgaaactaga 1260

ctcaatatga tcaataataa gattgatgac caaatacaag acgtatgggc atataatgca 1320

gaattgctag tactacttga aaatcaaaaa acactcgatg agcatgatgc gaacgtgaac 1380

aatctatata acaaggtgaa gagggcactg ggctccaatg ctatggaaga tgggaaaggc 1440

tgtttcgagc tataccataa atgtgatgat cagtgcatgg aaacaattcg gaacgggacc 1500

tataatagga gaaagtatag agaggaatca agactagaaa ggcagaaaat agagggggtt 1560

aagctggaat ctgagggaac ttacaaaatc ctcaccattt attcgactgt cgcctcatct 1620

cttgtgcttg caatggggtt tgctgccttc ctgttctggg ccatgtccaa tggatcttgc 1680

agatgcaaca tttgtatata a 1701

<210> 156
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF**-H9HK107399.S1-6r

<400> 156
actaaagaaa ataggccttt atatacaaat gttgcatctg caagatccat 50

<210> 157
<211> 566
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H9 Hong Kong

<400> 157

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Lys Ile Cys Ile Gly His Gln
20 25 30

Ser Thr Asn Ser Thr Glu Thr Val Asp Thr Leu Thr Glu Thr Asn Val
35 40 45

Pro Val Thr His Ala Lys Glu Leu Leu His Thr Glu His Asn Gly Met
50 55 60

Leu Cys Ala Thr Ser Leu Gly His Pro Leu Ile Leu Asp Thr Cys Thr
65 70 75 80

Ile Glu Gly Leu Val Tyr Gly Asn Pro Ser Cys Asp Leu Leu Leu Gly
85 90 95

Gly Arg Glu Trp Ser Tyr Ile Val Glu Arg Ser Ser Ala Val Asn Gly
100 105 110

Thr Cys Tyr Pro Gly Asn Val Glu Asn Leu Glu Glu Leu Arg Thr Leu
115 120 125

Phe Ser Ser Ala Ser Ser Tyr Gln Arg Ile Gln Ile Phe Pro Asp Thr
130 135 140

Thr Trp Asn Val Thr Tyr Thr Gly Thr Ser Arg Ala Cys Ser Gly Ser
145 150 155 160

```
Phe Tyr Arg Ser Met Arg Trp Leu Thr Gln Lys Ser Gly Phe Tyr Pro
                165                 170                 175

Val Gln Asp Ala Gln Tyr Thr Asn Asn Arg Gly Lys Ser Ile Leu Phe
            180                 185                 190

Val Trp Gly Ile His His Pro Pro Thr Tyr Thr Glu Gln Thr Asn Leu
        195                 200                 205

Tyr Ile Arg Asn Asp Thr Thr Thr Ser Val Thr Thr Glu Asp Leu Asn
    210                 215                 220

Arg Thr Phe Lys Pro Val Ile Gly Pro Arg Pro Leu Val Asn Gly Leu
225                 230                 235                 240

Gln Gly Arg Ile Asp Tyr Tyr Trp Ser Val Leu Lys Pro Gly Gln Thr
                245                 250                 255

Leu Arg Val Arg Ser Asn Gly Asn Leu Ile Ala Pro Trp Tyr Gly His
            260                 265                 270

Val Leu Ser Gly Gly Ser His Gly Arg Ile Leu Lys Thr Asp Leu Lys
        275                 280                 285

Gly Gly Asn Cys Val Val Gln Cys Gln Thr Glu Lys Gly Gly Leu Asn
    290                 295                 300

Ser Thr Leu Pro Phe His Asn Ile Ser Lys Tyr Ala Phe Gly Thr Cys
305                 310                 315                 320

Pro Lys Tyr Val Arg Val Asn Ser Leu Lys Leu Ala Val Gly Leu Arg
                325                 330                 335

Asn Val Pro Ala Arg Ser Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly
            340                 345                 350

Phe Ile Glu Gly Gly Trp Pro Gly Leu Val Ala Gly Trp Tyr Gly Phe
        355                 360                 365

Gln His Ser Asn Asp Gln Gly Val Gly Met Ala Ala Asp Arg Asp Ser
    370                 375                 380
```

```
Thr Gln Lys Ala Ile Asp Lys Ile Thr Ser Lys Val Asn Asn Ile Val
385                 390                 395                 400

Asp Lys Met Asn Lys Gln Tyr Glu Ile Ile Asp His Glu Phe Ser Glu
                405                 410                 415

Val Glu Thr Arg Leu Asn Met Ile Asn Asn Lys Ile Asp Asp Gln Ile
            420                 425                 430

Gln Asp Val Trp Ala Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn
        435                 440                 445

Gln Lys Thr Leu Asp Glu His Asp Ala Asn Val Asn Asn Leu Tyr Asn
    450                 455                 460

Lys Val Lys Arg Ala Leu Gly Ser Asn Ala Met Glu Asp Gly Lys Gly
465                 470                 475                 480

Cys Phe Glu Leu Tyr His Lys Cys Asp Asp Gln Cys Met Glu Thr Ile
                485                 490                 495

Arg Asn Gly Thr Tyr Asn Arg Arg Lys Tyr Arg Glu Glu Ser Arg Leu
            500                 505                 510

Glu Arg Gln Lys Ile Glu Gly Val Lys Leu Glu Ser Glu Gly Thr Tyr
        515                 520                 525

Lys Ile Leu Thr Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Ala
    530                 535                 540

Met Gly Phe Ala Ala Phe Leu Phe Trp Ala Met Ser Asn Gly Ser Cys
545                 550                 555                 560

Arg Cys Asn Ile Cys Ile
                565
```

<210> 158
<211> 1656
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/H9 Hong Kong với vòng phân giải protein loại bớt

<400> 158
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggataaaat ctgcatcggc caccagtcaa caaactccac agaaactgtg 120

gacacgctaa cagaaaccaa tgttcctgtg acacatgcca aagaattgct ccacacagag 180

cataatggaa tgctgtgtgc aacaagcctg ggacatcccc tcattctaga cacatgcact 240

attgaaggac tagtctatgg caacccttct tgtgacctgc tgttgggagg aagagaatgg 300

tcctacatcg tcgaaagatc atcagctgta aatggaacgt gttaccctgg gaatgtagaa 360

aacctagagg aactcaggac acttttttagt tccgctagtt cctaccaaag aatccaaatc 420

ttcccagaca caacctggaa tgtgacttac actggaacaa gcagagcatg ttcaggttca 480

ttctacagga gtatgagatg gctgactcaa aagagcggtt tttaccctgt tcaagacgcc 540

caatacacaa ataacagggg aaagagcatt cttttcgtgt ggggcataca tcacccaccc 600

acctataccg agcaaacaaa tttgtacata agaaacgaca caacaacaag cgtgacaaca 660

gaagatttga ataggacctt caaaccagtg atagggccaa ggccccttgt caatggtctg 720

cagggaagaa ttgattatta ttggtcggta ctaaaaccag gccaaacatt gcgagtacga 780

tccaatggga atctaattgc tccatggtat ggacacgttc tttcaggagg gagccatgga 840

agaatcctga agactgattt aaaaggtggt aattgtgtag tgcaatgtca gactgaaaaa 900

ggtggcttaa acagtacatt gccattccac aatatcagta aatatgcatt tggaacctgc 960

cccaaatatg taagagttaa tagtctcaaa ctggcagtcg gtctgaggaa cgtgcctggc 1020

ggaggttggc caggactagt cgctggctgg tatggtttcc agcattcaaa tgatcaaggg 1080

gttggtatgg ctgcagatag ggattcaact caaaaggcaa ttgataaaat aacatccaag 1140

gtgaataata tagtcgacaa gatgaacaag caatatgaaa taattgatca tgaatttagt 1200

gaggttgaaa ctagactcaa tatgatcaat aataagattg atgaccaaat acaagacgta 1260

tgggcatata atgcagaatt gctagtacta cttgaaaatc aaaaaacact cgatgagcat 1320

gatgcgaacg tgaacaatct atataacaag gtgaagaggg cactgggctc caatgctatg 1380

gaagatggga aaggctgttt cgagctatac cataaatgtg atgatcagtg catggaaaca 1440

attcggaacg ggacctataa taggagaaag tatagagagg aatcaagact agaaaggcag 1500

aaaatagagg gggttaagct ggaatctgag ggaacttaca aaatcctcac catttattcg 1560

actgtcgcct catctcttgt gcttgcaatg gggtttgctg ccttcctgtt ctgggccatg 1620

tccaatggat cttgcagatg caacatttgt atataa 1656

<210> 159
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H9HK107399(Prl-).r

<400> 159
gtcctggcca acctccgcca ggcacgttcc tcagaccgac tgccagtt 48

<210> 160
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H9HK107399(Prl-).c

<400> 160
ggaacgtgcc tggcggaggt tggccaggac tagtcgctgg ctggtatg 48

<210> 161
<211> 551
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/H9 Hong Kong với vòng phân giải protein loại bớt

<400> 161

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Lys Ile Cys Ile Gly His Gln
20 25 30

Ser Thr Asn Ser Thr Glu Thr Val Asp Thr Leu Thr Glu Thr Asn Val
35 40 45

```
Pro Val Thr His Ala Lys Glu Leu Leu His Thr Glu His Asn Gly Met
    50                  55                  60                  .

Leu Cys Ala Thr Ser Leu Gly His Pro Leu Ile Leu Asp Thr Cys Thr
65                  70                  75                  80

Ile Glu Gly Leu Val Tyr Gly Asn Pro Ser Cys Asp Leu Leu Leu Gly
                85                  90                  95

Gly Arg Glu Trp Ser Tyr Ile Val Glu Arg Ser Ser Ala Val Asn Gly
            100                 105                 110

Thr Cys Tyr Pro Gly Asn Val Glu Asn Leu Glu Glu Leu Arg Thr Leu
        115                 120                 125

Phe Ser Ser Ala Ser Ser Tyr Gln Arg Ile Gln Ile Phe Pro Asp Thr
    130                 135                 140

Thr Trp Asn Val Thr Tyr Thr Gly Thr Ser Arg Ala Cys Ser Gly Ser
145                 150                 155                 160

Phe Tyr Arg Ser Met Arg Trp Leu Thr Gln Lys Ser Gly Phe Tyr Pro
                165                 170                 175

Val Gln Asp Ala Gln Tyr Thr Asn Asn Arg Gly Lys Ser Ile Leu Phe
            180                 185                 190

Val Trp Gly Ile His His Pro Pro Thr Tyr Thr Glu Gln Thr Asn Leu
        195                 200                 205

Tyr Ile Arg Asn Asp Thr Thr Thr Ser Val Thr Thr Glu Asp Leu Asn
    210                 215                 220

Arg Thr Phe Lys Pro Val Ile Gly Pro Arg Pro Leu Val Asn Gly Leu
225                 230                 235                 240
                                                            .
Gln Gly Arg Ile Asp Tyr Tyr Trp Ser Val Leu Lys Pro Gly Gln Thr
                245                 250                 255
```

```
Leu Arg Val Arg Ser Asn Gly Asn Leu Ile Ala Pro Trp Tyr Gly His
            260                 265                 270

Val Leu Ser Gly Gly Ser His Gly Arg Ile Leu Lys Thr Asp Leu Lys
        275                 280                 285

Gly Gly Asn Cys Val Val Gln Cys Gln Thr Glu Lys Gly Gly Leu Asn
    290                 295                 300

Ser Thr Leu Pro Phe His Asn Ile Ser Lys Tyr Ala Phe Gly Thr Cys
305                 310                 315                 320

Pro Lys Tyr Val Arg Val Asn Ser Leu Lys Leu Ala Val Gly Leu Arg
                325                 330                 335

Asn Val Pro Gly Gly Gly Trp Pro Gly Leu Val Ala Gly Trp Tyr Gly
            340                 345                 350

Phe Gln His Ser Asn Asp Gln Gly Val Gly Met Ala Ala Asp Arg Asp
        355                 360                 365

Ser Thr Gln Lys Ala Ile Asp Lys Ile Thr Ser Lys Val Asn Asn Ile
    370                 375                 380

Val Asp Lys Met Asn Lys Gln Tyr Glu Ile Ile Asp His Glu Phe Ser
385                 390                 395                 400

Glu Val Glu Thr Arg Leu Asn Met Ile Asn Asn Lys Ile Asp Asp Gln
                405                 410                 415

Ile Gln Asp Val Trp Ala Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu
            420                 425                 430

Asn Gln Lys Thr Leu Asp Glu His Asp Ala Asn Val Asn Asn Leu Tyr
        435                 440                 445

Asn Lys Val Lys Arg Ala Leu Gly Ser Asn Ala Met Glu Asp Gly Lys
    450                 455                 460

Gly Cys Phe Glu Leu Tyr His Lys Cys Asp Asp Gln Cys Met Glu Thr
465                 470                 475                 480
```

Ile Arg Asn Gly Thr Tyr Asn Arg Arg Lys Tyr Arg Glu Glu Ser Arg
485 490 495

Leu Glu Arg Gln Lys Ile Glu Gly Val Lys Leu Glu Ser Glu Gly Thr
500 505 510

Tyr Lys Ile Leu Thr Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu
515 520 525

Ala Met Gly Phe Ala Ala Phe Leu Phe Trp Ala Met Ser Asn Gly Ser
530 535 540

Cys Arg Cys Asn Ile Cys Ile
545 550

<210> 162
<211> 1785
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Malaysia

<400> 162
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggatcgaat ctgcactggg ataacatcgt caaactcacc acatgttgtc 120

aaaactgcta ctcaagggga ggtcaatgtg actggtgtaa taccactgac aacaacaccc 180

accaaatctc attttgcaaa tctcaaagga acagaaacca gagggaaact atgcccaaaa 240

tgcctcaact gcacagatct ggacgtggcc ttgggcagac caaaatgcac ggggaacata 300

ccctcggcaa gagtttcaat actccatgaa gtcagacctg ttacatctgg gtgctttcct 360

ataatgcacg acagaacaaa aattagacag ctgcctaaac ttctcagagg atacgaacat 420

atcaggttat caactcataa cgttatcaat gcagaaaatg caccaggagg accctacaaa 480

attggaacct cagggtcttg ccctaacgtt accaatggaa acggattttt cgcaacaatg 540

gcttgggccg tcccaaaaaa cgacaacaac aaaacagcaa caaattcatt aacaatagaa 600

gtaccataca tttgtacaga aggagaagac caaattaccg tttgggggtt ccactctgat 660

aacgaaaccc aaatggcaaa gctctatggg gactcaaagc cccagaagtt cacctcatct 720

gccaacggag tgaccacaca ttacgtttca cagattggtg gcttcccaaa tcaaacagaa 780

gacggaggac taccacaaag cggtagaatt gttgttgatt acatggtgca aaaatctggg 840

aaaacaggaa caattaccta tcaaagaggt attttattgc ctcaaaaagt gtggtgcgca 900

agtggcagga gcaaggtaat aaaaggatcg ttgcctttaa ttggagaagc agattgcctc 960

cacgaaaaat acggtggatt aaacaaaagc aagccttact acacagggga acatgcaaag 1020

gccataggaa attgcccaat atgggtgaaa acacccttga agctggccaa tggaaccaaa 1080

tatagacctc ctgcaaaact attaaaggaa aggggtttct tcggagctat tgctggtttc 1140

ttagaaggag gatgggaagg aatgattgca ggttggcacg gatacacatc ccatggggca 1200

catggagtag cggtggcagc agaccttaag agcactcaag aggccataaa caagataaca 1260

aaaaatctca actctttgag tgagctggaa gtaaagaatc ttcaaagact aagcggtgcc 1320

atggatgaac tccacaacga aatactagaa ctagacgaga aagtggatga tctcagagct 1380

gatacaataa gctcacaaat agaactcgca gtcctgcttt ccaatgaagg aataataaac 1440

agtgaagatg agcatctctt ggcgcttgaa agaaagctga agaaaatgct gggcccctct 1500

gctgtagaga tagggaatgg atgctttgaa accaaacaca agtgcaacca gacctgtctc 1560

gacagaatag ctgctggtac ctttgatgca ggagaatttt ctctccccac ttttgattca 1620

ctgaatatta ctgctgcatc tttaaatgac gatggattgg ataatcatac tatactgctt 1680

tactactcaa ctgctgcctc cagtttggct gtaacattga tgatagctat ctttgttgtt 1740

tatatggtct ccagagacaa tgtttcttgc tccatctgtc tataa 1785

<210> 163
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF**-HBM250604.S1-6r

<400> 163
actaaagaaa ataggccttt atagacagat ggagcaagaa acattgtctc 50

<210> 164
<211> 594
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Malaysia

<400> 164

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
50 55 60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
85 90 95

Thr Gly Asn Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
115 120 125

Arg Gln Leu Pro Lys Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
130 135 140

Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
145 150 155 160

Ile Gly Thr Ser Gly Ser Cys Pro Asn Val Thr Asn Gly Asn Gly Phe
165 170 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Asn Asn Lys Thr
180 185 190

```
Ala Thr Asn Ser Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly
        195                 200                 205

Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu Thr Gln
    210                 215                 220

Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser
225                 230                 235                 240

Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro
                245                 250                 255

Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val
            260                 265                 270

Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln
        275                 280                 285

Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser
    290                 295                 300

Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu
305                 310                 315                 320

His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly
                325                 330                 335

Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro
            340                 345                 350

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys Leu Leu
        355                 360                 365

Lys Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu Gly Gly
    370                 375                 380

Trp Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala
385                 390                 395                 400

His Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile
                405                 410                 415
```

Asn Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys
420 425 430

Asn Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile
435 440 445

Leu Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser
450 455 460

Ser Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn
465 470 475 480

Ser Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met
485 490 495

Leu Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys
500 505 510

His Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe
515 520 525

Asp Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr
530 535 540

Ala Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu
545 550 555 560

Tyr Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Ile Ala
565 570 575

Ile Phe Val Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile
580 585 590

Cys Leu

<210> 165
<211> 1734
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Malaysia với vòng phân giải protein loại bớt

<400> 165
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggatcgaat ctgcactggg ataacatcgt caaactcacc acatgttgtc 120

aaaactgcta ctcaagggga ggtcaatgtg actggtgtaa taccactgac aacaacaccc 180

accaaatctc attttgcaaa tctcaaagga acagaaacca gagggaaact atgcccaaaa 240

tgcctcaact gcacagatct ggacgtggcc ttgggcagac caaaatgcac ggggaacata 300

ccctcggcaa gagtttcaat actccatgaa gtcagacctg ttacatctgg gtgctttcct 360

ataatgcacg acagaacaaa aattagacag ctgcctaaac ttctcagagg atacgaacat 420

atcaggttat caactcataa cgttatcaat gcagaaaatg caccaggagg accctacaaa 480

attggaacct cagggtcttg ccctaacgtt accaatggaa acggattttt cgcaacaatg 540

gcttgggccg tcccaaaaaa cgacaacaac aaaacagcaa caaattcatt aacaatagaa 600

gtaccataca tttgtacaga aggagaagac caaattaccg tttgggggtt ccactctgat 660

aacgaaaccc aaatggcaaa gctctatggg gactcaaagc cccagaagtt cacctcatct 720

gccaacggag tgaccacaca ttacgtttca cagattggtg gcttcccaaa tcaaacagaa 780

gacggaggac taccacaaag cggtagaatt gttgttgatt acatggtgca aaaatctggg 840

aaaacaggaa caattaccta tcaaagaggt attttattgc ctcaaaaagt gtggtgcgca 900

agtggcagga gcaaggtaat aaaaggatcg ttgcctttaa ttggagaagc agattgcctc 960

cacgaaaaat acggtggatt aaacaaaagc aagccttact acacagggga acatgcaaag 1020

gccataggaa attgcccaat atgggtgaaa acacccttga agctggccaa tggaaccaaa 1080

tatagacctc ctggtggagg atgggaagga atgattgcag gttggcacgg atacacatcc 1140

catggggcac atggagtagc ggtggcagca gaccttaaga gcactcaaga ggccataaac 1200

aagataacaa aaaatctcaa ctctttgagt gagctggaag taaagaatct tcaaagacta 1260

agcggtgcca tggatgaact ccacaacgaa atactagaac tagacgagaa agtggatgat 1320

ctcagagctg atacaataag ctcacaaata gaactcgcag tcctgctttc caatgaagga 1380

ataataaaca gtgaagatga gcatctcttg gcgcttgaaa gaaagctgaa gaaaatgctg 1440

ggcccctctg ctgtagagat agggaatgga tgctttgaaa ccaaacacaa gtgcaaccag 1500

acctgtctcg acagaatagc tgctggtacc tttgatgcag gagaattttc tctccccact 1560

tttgattcac tgaatattac tgctgcatct ttaaatgacg atggattgga taatcatact 1620

atactgcttt actactcaac tgctgcctcc agtttggctg taacattgat gatagctatc 1680

tttgttgttt atatggtctc cagagacaat gtttcttgct ccatctgtct ataa . 1734

<210> 166
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> HBM250604(PrL-).r

<400> 166
cattccttcc catcctccac caggaggtct atatttggtt ccattggc 48

<210> 167
<211> 48
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> HBM250604(PrL-).c

<400> 167
agacctcctg gtggaggatg ggaaggaatg attgcaggtt ggcacgga 48

<210> 168
<211> 577
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Malaysia với vòng phân giải protein loại bớt

<400> 168

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

```
Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
        35                  40                  45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser His
    50                  55                  60

Phe Ala Asn Leu Lys Gly Thr Glu Thr Arg Gly Lys Leu Cys Pro Lys
65                  70                  75                  80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Lys Cys
                85                  90                  95

Thr Gly Asn Ile Pro Ser Ala Arg Val Ser Ile Leu His Glu Val Arg
            100                 105                 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
        115                 120                 125

Arg Gln Leu Pro Lys Leu Leu Arg Gly Tyr Glu His Ile Arg Leu Ser
    130                 135                 140

Thr His Asn Val Ile Asn Ala Glu Asn Ala Pro Gly Gly Pro Tyr Lys
145                 150                 155                 160

Ile Gly Thr Ser Gly Ser Cys Pro Asn Val Thr Asn Gly Asn Gly Phe
                165                 170                 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asn Asp Asn Asn Lys Thr
            180                 185                 190

Ala Thr Asn Ser Leu Thr Ile Glu Val Pro Tyr Ile Cys Thr Glu Gly
        195                 200                 205

Glu Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Glu Thr Gln
    210                 215                 220

Met Ala Lys Leu Tyr Gly Asp Ser Lys Pro Gln Lys Phe Thr Ser Ser
225                 230                 235                 240

Ala Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro
                245                 250                 255
```

Asn Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val
260 265 270

Asp Tyr Met Val Gln Lys Ser Gly Lys Thr Gly Thr Ile Thr Tyr Gln
275 280 285

Arg Gly Ile Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser
290 295 300

Lys Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu
305 310 315 320

His Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly
325 330 335

Glu His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro
340 345 350

Leu Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp
355 360 365

Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His
370 375 380

Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn
385 390 395 400

Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn
405 410 415

Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu
420 425 430

Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser
435 440 445

Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser
450 455 460

Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu
465 470 475 480

Gly Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys His
485 490 495

Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asp
500 505 510

Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala
515 520 525

Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu Tyr
530 535 540

Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Ile Ala Ile
545 550 555 560

Phe Val Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile Cys
565 570 575

Leu

<210> 169
<211> 1782
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Massachusetts

<400> 169
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg ccgatcgaat ctgcactggg ataacatctt caaactcacc tcatgtggtc 120

aaaacagcta ctcaagggga ggtcaatgtg actggtgtga taccactaac aacaacacca 180

acaaaatctt attttgcaaa tctcaaagga acaaagacca gagggaaact atgcccagac 240

tgtctcaact gtacagatct ggatgtggcc ctgggcaggc caatgtgtgt gggaactaca 300

ccttctgcga aagcttcaat acttcacgaa gtcagacctg ttacatccgg gtgcttccct 360

ataatgcacg acagaacaaa aatcaggcaa ctagccaatc ttctcagagg atatgaaaat 420

atcaggttat caacccaaaa cgttatcgat gcagaaaagg caccaggagg accctacaga 480

cttggaacct caggatcttg ccctaacgct accagtaaaa gcggattttt cgcaacaatg 540

gcttgggctg tcccaaagga caacaacaaa aatgcaacga acccattaac agtagaagta 600

ccatacattt gtgcagaagg ggaagaccaa attactgttt gggggttcca ttcagataac 660

aaaacccaaa tgaagaacct ctatggagac tcaaatcctc aaaagttcac ctcatctgct 720

aatggagtaa ccacacatta tgtttctcag attggcggct tcccagatca aacagaagac 780

ggaggactac cacaaagcgg cagaattgtc gttgattaca tgatgcaaaa acctgggaaa 840

acaggaacaa ttgtctatca aagaggtgtt ttgttgcctc aaaaggtgtg gtgcgcgagt 900

ggcaggagca aagtaataaa agggtccttg cctttaattg gtgaagcaga ttgccttcat 960

gaaaaatacg gtggattaaa caaaagcaag ccttactaca caggagaaca tgcaaaagcc 1020

ataggaaatt gcccaatatg ggtgaaaaca cctttgaagc ttgccaatgg aaccaaatat 1080

agacctcctg caaaactatt aaaggaaagg ggtttcttcg gagctattgc tggtttccta 1140

gaaggaggat gggaaggaat gattgcaggt tggcacggat acacatctca cggagcacat 1200

ggagtggcag ttgctgcaga ccttaagagc acacaagaag ctataaacaa gataacaaaa 1260

aatctcaact ctttgagtga gctagaagta aagaatcttc aaaggctaag tggtgccatg 1320

gatgaactcc acaacgaaat actcgagctg gatgagaaag tggatgacct cagagctgac 1380

actataagtt cacaaataga acttgcagtc ttgctttcca acgaaggaat aataaacagt 1440

gaagacgagc atctattggc acttgagaga aaactaaaga aaatgctggg tccctctgct 1500

gtagacatag gaaatggatg cttcgaaacc aaacacaaat gcaaccagac ctgcttagac 1560

aggatagctg ctggcacctt taatgcagga gagttttctc tccccacttt tgattcattg 1620

aacattactg ctgcatcttt aaatgatgat ggattggata accatactat actgctctat 1680

tactcaactg ctgcttctag tttggctgta acattgatgc tagctatttt tattgtttat 1740

atggtctcca gagacaacgt ttcatgctcc atctgtctat aa 1782

<210> 170
<211> 593
<212> PRT
<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của PDISP/HA B Massachusetts

<400> 170

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser Tyr
50 55 60

Phe Ala Asn Leu Lys Gly Thr Lys Thr Arg Gly Lys Leu Cys Pro Asp
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Met Cys
85 90 95

Val Gly Thr Thr Pro Ser Ala Lys Ala Ser Ile Leu His Glu Val Arg
100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
115 120 125

Arg Gln Leu Ala Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser
130 135 140

Thr Gln Asn Val Ile Asp Ala Glu Lys Ala Pro Gly Gly Pro Tyr Arg
145 150 155 160

Leu Gly Thr Ser Gly Ser Cys Pro Asn Ala Thr Ser Lys Ser Gly Phe
165 170 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asp Asn Asn Lys Asn Ala
180 185 190

Thr Asn Pro Leu Thr Val Glu Val Pro Tyr Ile Cys Ala Glu Gly Glu
195 200 205

```
Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Lys Thr Gln Met
    210                 215                 220

Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala
225                 230                 235                 240

Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Gly Phe Pro Asp
                245                 250                 255

Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp
            260                 265                 270

Tyr Met Met Gln Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg
        275                 280                 285

Gly Val Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys
    290                 295                 300

Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu His
305                 310                 315                 320

Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly Glu
                325                 330                 335

His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro Leu
            340                 345                 350

Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Ala Lys Leu Leu Lys
        355                 360                 365

Glu Arg Gly Phe Phe Gly Ala Ile Ala Gly Phe Leu Glu Gly Gly Trp
    370                 375                 380

Glu Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His
385                 390                 395                 400

Gly Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn
                405                 410                 415
```

Lys Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn
420 425 430

Leu Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu
435 440 445

Glu Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser
450 455 460

Gln Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser
465 470 475 480

Glu Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu
485 490 495

Gly Pro Ser Ala Val Asp Ile Gly Asn Gly Cys Phe Glu Thr Lys His
500 505 510

Lys Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asn
515 520 525

Ala Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala
530 535 540

Ala Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu Tyr
545 550 555 560

Tyr Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Leu Ala Ile
565 570 575

Phe Ile Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile Cys
580 585 590

Leu

<210> 171
<211> 63
<212> ADN
<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit của peptit tín hiệu patatin A

<400> 171
atggcaacta ctaaaacttt tttaatttta ttttttatga tattagcaac tactagttca 60

aca 63

<210> 172
<211> 23
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của peptit tín hiệu patatin A

<400> 172

Met Ala Thr Thr Lys Thr Phe Leu Ile Leu Phe Phe Met Ile Leu Ala
1 5 10 15

Thr Thr Ser Ser Thr Cys Ala
20

<210> 173
<211> 517
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự CMPV HT

<400> 173
tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc 60

ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc ttgcgtgagc 120

gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgttttctt tcactgaagc 180

gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg tgtacttgtc 240

ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct gttcagcccc 300

atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct acttctgctt 360

gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt tctataagaa 420

atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga gaaagattgt 480

taagcttctg tatattctgc ccaaatttgt cgggccc 517

<210> 174

<400> 174
000

<210> 175

<400> 175
000

<210> 176

<400> 176
000

<210> 177

<400> 177
000

<210> 178

<400> 178
000

<210> 179

<400> 179
000

<210> 180

<400> 180
000

<210> 181

<400> 181
000

<210> 182

<400> 182
000

<210> 183

<400> 183
000

<210> 184

<400> 184
000

<210> 185

<400> 185
000

<210> 186

<400> 186
000

<210> 187
<211> 528
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> [00173] CPMV HT+ với trình tự thống nhất kozak của thực vật

<400> 187
tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc 60

ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc ttgcgtgagc 120

gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgttttctt tcactgaagc 180

gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg tgtacttgtc 240

ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct gttcagcccc 300

atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct acttctgctt 360

gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt tctataagaa 420

atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga gaaagattgt 480

taagcttctg tatattctgc ccaaatttgt tcgggcccaa taccgcgg 528

<210> 188
<211> 511
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của CPMV HT+ 511

<400> 188
tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc 60

ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc ttgcgtgagc 120

gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgttttctt tcactgaagc 180

gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg tgtacttgtc 240

ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct gttcagcccc 300

atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct acttctgctt 360

gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt tctataagaa 420

atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga gaaagattgt 480

taagcttctg tatattctgc ccaaatttga a 511

<210> 189
<211> 534
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của CPMV HT+ [WT115]

<400> 189
tattaaaatc ttaataggtt ttgataaaag cgaacgtggg gaaacccgaa ccaaaccttc 60

ttctaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc ttgcatgagc 120

gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgttttctt tcactgaagc 180

gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg tgtacttgtc 240

ctattcttgt cggtgtggtc ttgggaaaag aaagcttgct ggaggctgct gttcagcccc 300

atacattact tgttacgatt ctgctgactt tcggcgggtg caatatctct acttctgctt 360

gacgaggtat tgttgcctgt acttctttct tcttcttctt gctgattggt tctataagaa 420

atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga gaaagattgt 480

taagcttctg tatattctgc ccaaatttgt tcgggcccaa taccgcggag aaaa 534

<210> 190
<211> 56
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> HBF406(PrL-).r (cấu trúc 2102 and 2104)

<400> 190
tccttcccat cctccaccag gaggtctata tttggttcca ttggcgagct tcaaag 56

<210> 191

<211> 52
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> HBF406(PrL-).c (cấu trúc 2102 và 2104)

<400> 191
atatagacct cctggtggag gatgggaagg aatgattgca ggctggcacg ga 52

<210> 192
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF*-HBF406.s1-6r (cấu trúc 2102 và 2104)

<400> 192
actaaagaaa ataggccttt atagacagat ggagcatgaa acgttgtctc 50

<210> 193
<211> 1731
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Florida với vòng phân giải protein loại bớt (cấu trúc 2102 và 2104)

<400> 193
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggatcgaat ctgcactgga ataacatctt caaactcacc tcatgtggtc 120

aaaacagcca ctcaagggga ggtcaatgtg actggtgtga taccactaac aacaacacca 180

acaaaatctt attttgcaaa tctcaaagga acaaggacca gagggaaact atgcccagac 240

tgtctcaact gcacagatct ggatgtggct ttgggcagac caatgtgtgt ggggaccaca 300

ccttcggcga aggcttcaat actccacgaa gtcaaacctg ttacatccgg gtgctttcct 360

ataatgcacg acagaacaaa aatcaggcaa ctacccaatc ttctcagagg atatgaaaat 420

atcaggctat caacccaaaa cgtcatcgat gcggaaaagg caccaggagg accctacaga 480

cttggaacct caggatcttg ccctaacgct accagtaaga gcggattttt cgcaacaatg 540

gcttgggctg tcccaaagga caacaacaaa aatgcaacga acccactaac agtagaagta 600

ccatacattt gtacagaagg ggaagaccaa atcactgttt gggggttcca ttcagataac 660

aaaacccaaa tgaagaacct ctatggagac tcaaatcctc aaaagttcac ctcatctgct 720

aatggagtaa ccacacacta tgtttctcag attggcagct tcccagatca aacagaagac 780

ggaggactac cacaaagcgg caggattgtt gttgattaca tgatgcaaaa acctgggaaa 840

acaggaacaa ttgtctacca aagaggtgtt ttgttgcctc aaaaggtgtg gtgcgcgagt 900

ggcaggagca aagtaataaa agggtccttg cctttaattg gtgaagcaga ttgccttcat 960

gaaaaatacg gtggattaaa caaaagcaag ccttactaca caggagaaca tgcaaaagcc 1020

ataggaaatt gcccaatatg ggtgaaaaca cctttgaagc tcgccaatgg aaccaaatat 1080

agacctcctg gtggaggatg ggaaggaatg attgcaggct ggcacggata cacatctcac 1140

ggagcacatg gagtggcagt ggcggcggac cttaagagta cgcaagaagc tataaacaag 1200

ataacaaaaa atctcaattc tttgagtgag ctagaagtaa agaatcttca aagactaagt 1260

ggtgccatgg atgaactcca caacgaaata ctcgagctgg atgagaaagt ggatgatctc 1320

agagctgaca ctataagctc gcaaatagaa cttgcagtct tgctttccaa cgaaggaata 1380

ataaacagtg aagatgagca tctattggca cttgagagaa aactaaagaa aatgctgggt 1440

ccctctgctg tagagatagg aaatggatgc ttcgaaacca aacacaagtg caaccagacc 1500

tgcttagaca ggatagctgc tggcaccttt aatgcaggag aattttctct ccccactttt 1560

gattcactga acattactgc tgcatcttta aatgatgatg gattggataa ccatactata 1620

ctgctctatt actcaactgc tgcttctagt ttggctgtaa cattgatgct agctattttt 1680

attgtttata tggtctccag agacaacgtt tcatgctcca tctgtctata a 1731

<210> 194
<211> 576
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Florida với vòng phân giải protein loại bớt (cấu trúc 2102 và 2104)

<400> 194

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr

20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser Tyr
50 55 60

Phe Ala Asn Leu Lys Gly Thr Arg Thr Arg Gly Lys Leu Cys Pro Asp
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Met Cys
85 90 95

Val Gly Thr Thr Pro Ser Ala Lys Ala Ser Ile Leu His Glu Val Lys
100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
115 120 125

Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser
130 135 140

Thr Gln Asn Val Ile Asp Ala Glu Lys Ala Pro Gly Gly Pro Tyr Arg
145 150 155 160

Leu Gly Thr Ser Gly Ser Cys Pro Asn Ala Thr Ser Lys Ser Gly Phe
165 170 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asp Asn Asn Lys Asn Ala
180 185 190

Thr Asn Pro Leu Thr Val Glu Val Pro Tyr Ile Cys Thr Glu Gly Glu
195 200 205

Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Lys Thr Gln Met
210 215 220

Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala
225 230 235 240

```
Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Ser Phe Pro Asp
                245                 250                 255

Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp
            260                 265                 270

Tyr Met Met Gln Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg
        275                 280                 285

Gly Val Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys
    290                 295                 300

Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu His
305                 310                 315                 320

Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly Glu
                325                 330                 335

His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro Leu
            340                 345                 350

Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp Glu
        355                 360                 365

Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His Gly
    370                 375                 380

Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn Lys
385                 390                 395                 400

Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn Leu
                405                 410                 415

Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu Glu
            420                 425                 430

Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser Gln
        435                 440                 445

Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser Glu
```

450 455 460

Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu Gly
465 470 475 480

Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys His Lys
485 490 495

Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asn Ala
500 505 510

Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala Ala
515 520 525

Ser Leu Asn Asp Asp Gly Leu Asp Asn His Thr Ile Leu Leu Tyr Tyr
530 535 540

Ser Thr Ala Ala Ser Ser Leu Ala Val Thr Leu Met Leu Ala Ile Phe
545 550 555 560

Ile Val Tyr Met Val Ser Arg Asp Asn Val Ser Cys Ser Ile Cys Leu
565 570 575

<210> 195
<211> 3486
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 2102

<400> 195
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta ataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccaatacc gcggagaaaa tggcgaaaaa cgttgcgatt ttcggcttat tgttttctct 1320

tcttgtgttg gttccttctc agatcttcgc ggatcgaatc tgcactggaa taacatcttc 1380

aaactcacct catgtggtca aaacagccac tcaaggggag gtcaatgtga ctggtgtgat 1440

accactaaca acaacaccaa caaaatctta ttttgcaaat ctcaaaggaa caaggaccag 1500

agggaaacta tgcccagact gtctcaactg cacagatctg gatgtggctt tgggcagacc 1560

aatgtgtgtg gggaccacac cttcggcgaa ggcttcaata ctccacgaag tcaaacctgt 1620

tacatccggg tgctttccta taatgcacga cagaacaaaa atcaggcaac tacccaatct 1680

tctcagagga tatgaaaata tcaggctatc aacccaaaac gtcatcgatg cggaaaaggc 1740

accaggagga ccctacagac ttggaacctc aggatcttgc cctaacgcta ccagtaagag 1800

cggatttttc gcaacaatgg cttgggctgt cccaaaggac aacaacaaaa atgcaacgaa 1860

cccactaaca gtagaagtac catacatttg tacagaaggg gaagaccaaa tcactgtttg 1920

ggggttccat tcagataaca aaacccaaat gaagaacctc tatggagact caaatcctca 1980

aaagttcacc tcatctgcta atggagtaac cacacactat gtttctcaga ttggcagctt 2040

cccagatcaa acagaagacg gaggactacc acaaagcggc aggattgttg ttgattacat 2100

gatgcaaaaa cctgggaaaa caggaacaat tgtctaccaa agaggtgttt tgttgcctca 2160

aaaggtgtgg tgcgcgagtg gcaggagcaa agtaataaaa gggtccttgc ctttaattgg 2220

tgaagcagat tgccttcatg aaaaatacgg tggattaaac aaaagcaagc cttactacac 2280

aggagaacat gcaaaagcca taggaaattg cccaatatgg gtgaaaacac ctttgaagct 2340

cgccaatgga accaaatata gacctcctgg tggaggatgg gaaggaatga ttgcaggctg 2400

gcacggatac acatctcacg gagcacatgg agtggcagtg gcggcggacc ttaagagtac 2460

gcaagaagct ataaacaaga taacaaaaaa tctcaattct ttgagtgagc tagaagtaaa 2520

gaatcttcaa agactaagtg gtgccatgga tgaactccac aacgaaatac tcgagctgga 2580

tgagaaagtg gatgatctca gagctgacac tataagctcg caaatagaac ttgcagtctt 2640

gctttccaac gaaggaataa taaacagtga agatgagcat ctattggcac ttgagagaaa 2700

actaaagaaa atgctgggtc cctctgctgt agagatagga aatggatgct tcgaaaccaa 2760

acacaagtgc aaccagacct gcttagacag gatagctgct ggcaccttta atgcaggaga 2820

attttctctc cccacttttg attcactgaa cattactgct gcatctttaa atgatgatgg 2880

attggataac catactatac tgctctatta ctcaactgct gcttctagtt tggctgtaac 2940

attgatgcta gctattttta ttgtttatat ggtctccaga gacaacgttt catgctccat 3000

ctgtctataa aggcctattt tctttagttt gaatttactg ttattcggtg tgcatttcta 3060

tgtttggtga gcggttttct gtgctcagag tgtgtttatt ttatgtaatt taatttcttt 3120

gtgagctcct gtttagcagg tcgtcccttc agcaaggaca caaaaagatt ttaattttat 3180

taaaaaaaaa aaaaaaaaag accgggaatt cgatatcaag cttatcgacc tgcagatcgt 3240

tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt 3300

atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg 3360

ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta atacgcgata 3420

gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta 3480

ctagat 3486

<210> 196
<211> 5368
<212> ADN

<213> Trình tự nhân tạo

<220>
<223> Cátxét 2104

<400> 196
tgttgttgtg actccgaggg gttgcctcaa actctatctt ataaccggcg tggaggcatg 60

gaggcagggg tattttggtc attttaatag atagtggaaa atgacgtgga atttacttaa 120

agacgaagtc tttgcgacaa ggggggggccc acgccgaatt taatattacc ggcgtggccc 180

ccccttatcg cgagtgcttt agcacgagcg gtccagattt aaagtagaaa atttcccgcc 240

cactagggtt aaaggtgttc acactataaa agcatatacg atgtgatggt atttggtcga 300

caagcttgca tgccggtcaa catggtggag cacgacacac ttgtctactc caaaaatatc 360

aaagatacag tctcagaaga ccaaagggca attgagactt ttcaacaaag ggtaatatcc 420

ggaaacctcc tcggattcca ttgcccagct atctgtcact ttattgtgaa gatagtggaa 480

aaggaaggtg gctcctacaa atgccatcat tgcgataaag gaaaggccat cgttgaagat 540

gcctctgccg acagtggtcc caaagatgga cccccaccca cgaggagcat cgtggaaaaa 600

gaagacgttc caaccacgtc ttcaaagcaa gtggattgat gtgataacat ggtggagcac 660

gacacacttg tctactccaa aaatatcaaa gatacagtct cagaagacca aagggcaatt 720

gagacttttc aacaaagggt aatatccgga aacctcctcg gattccattg cccagctatc 780

tgtcacttta ttgtgaagat agtggaaaag gaaggtggct cctacaaatg ccatcattgc 840

gataaaggaa aggccatcgt tgaagatgcc tctgccgaca gtggtcccaa agatggaccc 900

ccacccacga ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc aaagcaagtg 960

gattgatgtg atatctccac tgacgtaagg gatgacgcac aatcccacta tccttcgcaa 1020

gacccttcct ctatataagg aagttcattt catttggaga ggtattaaaa tcttaatagg 1080

ttttgataaa agcgaacgtg gggaaacccg aaccaaacct tcttctaaac tctctctcat 1140

ctctcttaaa gcaaacttct ctcttgtctt tcttgcgtga gcgatcttca acgttgtcag 1200

atcgtgcttc ggcaccagta caacgttttc tttcactgaa gcgaaatcaa agatctcttt 1260

gtggacacgt agtgcggcgc cattaaataa cgtgtacttg tcctattctt gtcggtgtgg 1320

tcttgggaaa agaaagcttg ctggaggctg ctgttcagcc ccatacatta cttgttacga 1380

ttctgctgac tttcggcggg tgcaatatct ctacttctgc ttgacgaggt attgttgcct 1440

gtacttcttt cttcttcttc ttgctgattg gttctataag aaatctagta ttttctttga 1500

aacagagttt tcccgtggtt ttcgaacttg gagaaagatt gttaagcttc tgtatattct 1560

gcccaaattt gtcgggccca ataccgcgga gaaaatggcg aaaaacgttg cgattttcgg 1620

cttattgttt tctcttcttg tgttggttcc ttctcagatc ttcgcggatc gaatctgcac 1680

tggaataaca tcttcaaact cacctcatgt ggtcaaaaca gccactcaag gggaggtcaa 1740

tgtgactggt gtgataccac taacaacaac accaacaaaa tcttattttg caaatctcaa 1800

aggaacaagg accagaggga aactatgccc agactgtctc aactgcacag atctggatgt 1860

ggctttgggc agaccaatgt gtgtggggac cacaccttcg gcgaaggctt caatactcca 1920

cgaagtcaaa cctgttacat ccgggtgctt tcctataatg cacgacagaa caaaaatcag 1980

gcaactaccc aatcttctca gaggatatga aaatatcagg ctatcaaccc aaaacgtcat 2040

cgatgcggaa aaggcaccag gaggacccta cagacttgga acctcaggat cttgccctaa 2100

cgctaccagt aagagcggat ttttcgcaac aatggcttgg gctgtcccaa aggacaacaa 2160

caaaaatgca acgaacccac taacagtaga agtaccatac atttgtacag aaggggaaga 2220

ccaaatcact gtttgggggt tccattcaga taacaaaacc caaatgaaga acctctatgg 2280

agactcaaat cctcaaaagt tcacctcatc tgctaatgga gtaaccacac actatgtttc 2340

tcagattggc agcttcccag atcaaacaga agacggagga ctaccacaaa gcggcaggat 2400

tgttgttgat tacatgatgc aaaaacctgg gaaaacagga acaattgtct accaaagagg 2460

tgttttgttg cctcaaaagg tgtggtgcgc gagtggcagg agcaaagtaa taaaagggtc 2520

cttgccttta attggtgaag cagattgcct tcatgaaaaa tacggtggat taaacaaaag 2580

caagccttac tacacaggag aacatgcaaa agccatagga aattgcccaa tatgggtgaa 2640

aacacctttg aagctcgcca atggaaccaa atatagacct cctggtggag gatgggaagg 2700

aatgattgca ggctggcacg gatacacatc tcacggagca catggagtgg cagtggcggc 2760

ggaccttaag agtacgcaag aagctataaa caagataaca aaaaatctca attctttgag 2820

tgagctagaa gtaaagaatc ttcaaagact aagtggtgcc atggatgaac tccacaacga 2880

aatactcgag ctggatgaga aagtggatga tctcagagct gacactataa gctcgcaaat 2940

agaacttgca gtcttgcttt ccaacgaagg aataataaac agtgaagatg agcatctatt 3000

ggcacttgag agaaaactaa agaaaatgct gggtccctct gctgtagaga taggaaatgg 3060

atgcttcgaa accaaacaca agtgcaacca gacctgctta gacaggatag ctgctggcac 3120

ctttaatgca ggagaatttt ctctccccac ttttgattca ctgaacatta ctgctgcatc 3180

tttaaatgat gatggattgg ataaccatac tatactgctc tattactcaa ctgctgcttc 3240

tagtttggct gtaacattga tgctagctat ttttattgtt tatatggtct ccagagacaa 3300

cgtttcatgc tccatctgtc tataaaggcc tattttcttt agtttgaatt tactgttatt 3360

cggtgtgcat ttctatgttt ggtgagcggt tttctgtgct cagagtgtgt ttattttatg 3420

taatttaatt tctttgtgag ctcctgttta gcaggtcgtc ccttcagcaa ggacacaaaa 3480

agattttaat tttattaaaa aaaaaaaaaa aaaagaccgg gaattcgata tcaagcttat 3540

cgacctgcag atcgttcaaa catttggcaa taaagtttct taagattgaa tcctgttgcc 3600

ggtcttgcga tgattatcat ataatttctg ttgaattacg ttaagcatgt aataattaac 3660

atgtaatgca tgacgttatt tatgagatgg gtttttatga ttagagtccc gcaattatac 3720

atttaatacg cgatagaaaa caaaatatag cgcgcaaact aggataaatt atcgcgcgcg 3780

gtgtcatcta tgttactaga tctctagagt ctcaagcttg gcgcgccata aaatgattat 3840

tttatgaata tatttcattg tgcaagtaga tagaaattac atatgttaca taacacacga 3900

aataaacaaa aaaagacaat ccaaaaacaa acaccccaaa aaaaataatc actttagata 3960

aactcgtatg aggagaggca cgttcagtga ctcgacgatt cccgagcaaa aaaagtctcc 4020

ccgtcacaca tatagtgggt gacgcaatta tctttaaagt aatccttctg ttgacttgtc 4080

attgataaca tccagtcttc gtcaggattg caaagaatta tagaagggat cccacctttt 4140

attttcttct tttttccata tttagggttg acagtgaaat cagactggca acctattaat 4200

tgcttccaca atgggacgaa cttgaagggg atgtcgtcga tgatattata ggtggcgtgt 4260

tcatcgtagt tggtgaaatc gatggtaccg ttccaatagt tgtgtcgtcc gagacttcta 4320

gcccaggtgg tctttccggt acgagttggt ccgcagatgt agaggctggg gtgtcggatt 4380

ccattccttc cattgtcctg gttaaatcgg ccatccattc aaggtcagat tgagcttgtt 4440

ggtatgagac aggatgtatg taagtataag cgtctatgct tacatggtat agatgggttt 4500

ccctccagga gtgtagatct tcgtggcagc gaagatctga ttctgtgaag ggcgacacat 4560

acggttcagg ttgtggaggg aataatttgt tggctgaata ttccagccat tgaagttttg 4620

ttgcccattc atgagggaat tcttccttga tcatgtcaag atattcctcc ttagacgttg 4680

cagtctggat aatagttctc catcgtgcgt cagatttgcg aggagagacc ttatgatctc 4740

ggaaatctcc tctggtttta atatctccgt cctttgatat gtaatcaagg acttgtttag 4800

agtttctagc tggctggata ttagggtgat ttccttcaaa atcgaaaaaa gaaggatccc 4860

taatacaagg ttttttatca agctggagaa gagcatgata gtgggtagtg ccatcttgat 4920

gaagctcaga agcaacacca aggaagaaaa taagaaaagg tgtgagtttc tcccagagaa 4980

actggaataa atcatctctt tgagatgagc acttgggata ggtaaggaaa acatatttag 5040

attggagtct gaagttctta ctagcagaag gcatgttgtt gtgactccga ggggttgcct 5100

caaactctat cttataaccg gcgtggaggc atggaggcag gggtattttg gtcattttaa 5160

tagatagtgg aaaatgacgt ggaatttact taaagacgaa gtctttgcga caaggggggg 5220

cccacgccga atttaatatt accggcgtgg ccccccctta tcgcgagtgc tttagcacga 5280

gcggtccaga tttaaagtag aaaatttccc gcccactagg gttaaaggtg ttcacactat 5340

aaaagcatat acgatgtgat ggtatttg 5368

<210> 197
<211> 46
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF-H1cTMCT.S1-4r

<400> 197
actaaagaaa ataggccttt aaatacatat tctacactgt agagac 46

<210> 198
<211> 1731
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự nucleotit của PDISP/HA B Florida+H1Cal TMCT với vòng phân giải protein loại bớt (cấu trúc 2106 và 2108)

<400> 198
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt ggttccttct 60

cagatcttcg cggatcgaat ctgcactgga ataacatctt caaactcacc tcatgtggtc 120

aaaacagcca ctcaagggga ggtcaatgtg actggtgtga taccactaac aacaacacca 180

acaaaatctt attttgcaaa tctcaaagga acaaggacca gagggaaact atgcccagac 240

tgtctcaact gcacagatct ggatgtggct ttgggcagac caatgtgtgt ggggaccaca 300

ccttcggcga aggcttcaat actccacgaa gtcaaacctg ttacatccgg gtgctttcct 360

ataatgcacg acagaacaaa aatcaggcaa ctacccaatc ttctcagagg atatgaaaat 420

atcaggctat caacccaaaa cgtcatcgat gcggaaaagg caccaggagg accctacaga 480

cttggaacct caggatcttg ccctaacgct accagtaaga gcggattttt cgcaacaatg 540

gcttgggctg tcccaaagga caacaacaaa aatgcaacga acccactaac agtagaagta 600

ccatacattt gtacagaagg ggaagaccaa atcactgttt gggggttcca ttcagataac 660

aaaacccaaa tgaagaacct ctatggagac tcaaatcctc aaaagttcac ctcatctgct 720

aatggagtaa ccacacacta tgtttctcag attggcagct tcccagatca aacagaagac 780

ggaggactac cacaaagcgg caggattgtt gttgattaca tgatgcaaaa acctgggaaa 840

acaggaacaa ttgtctacca aagaggtgtt ttgttgcctc aaaaggtgtg gtgcgcgagt 900

ggcaggagca aagtaataaa agggtccttg cctttaattg gtgaagcaga ttgccttcat 960

gaaaaatacg gtggattaaa caaaagcaag ccttactaca caggagaaca tgcaaaagcc 1020

ataggaaatt gcccaatatg ggtgaaaaca cctttgaagc tcgccaatgg aaccaaatat 1080

agacctcctg gtggaggatg ggaaggaatg attgcaggct ggcacggata cacatctcac 1140

ggagcacatg gagtggcagt ggcggcggac cttaagagta cgcaagaagc tataaacaag 1200

ataacaaaaa atctcaattc tttgagtgag ctagaagtaa agaatcttca aagactaagt 1260

ggtgccatgg atgaactcca caacgaaata ctcgagctgg atgagaaagt ggatgatctc 1320

agagctgaca ctataagctc gcaaatagaa cttgcagtct tgctttccaa cgaaggaata 1380

ataaacagtg aagatgagca tctattggca cttgagagaa aactaaagaa aatgctgggt 1440

ccctctgctg tagagatagg aaatggatgc ttcgaaacca aacacaagtg caaccagacc 1500

tgcttagaca ggatagctgc tggcaccttt aatgcaggag aattttctct ccccactttt 1560

gattcactga acattactgc tgcatcttta aatgatgatg gattggataa ctaccagatt 1620

ttggcgatct attcaactgt cgccagttca ttggtactgg tagtctccct gggggcaatc 1680

agtttctgga tgtgctctaa tgggtctcta cagtgtagaa tatgtattta a 1731

<210> 199
<211> 576

<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit amin của PDISP/HA B Florida+H1Cal TMCT với vòng phân giải protein loại bớt (cấu trúc 2106 và 2108)

<400> 199

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Arg Ile Cys Thr Gly Ile Thr
20 25 30

Ser Ser Asn Ser Pro His Val Val Lys Thr Ala Thr Gln Gly Glu Val
35 40 45

Asn Val Thr Gly Val Ile Pro Leu Thr Thr Thr Pro Thr Lys Ser Tyr
50 55 60

Phe Ala Asn Leu Lys Gly Thr Arg Thr Arg Gly Lys Leu Cys Pro Asp
65 70 75 80

Cys Leu Asn Cys Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Met Cys
85 90 95

Val Gly Thr Thr Pro Ser Ala Lys Ala Ser Ile Leu His Glu Val Lys
100 105 110

Pro Val Thr Ser Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile
115 120 125

Arg Gln Leu Pro Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser
130 135 140

Thr Gln Asn Val Ile Asp Ala Glu Lys Ala Pro Gly Gly Pro Tyr Arg
145 150 155 160

Leu Gly Thr Ser Gly Ser Cys Pro Asn Ala Thr Ser Lys Ser Gly Phe
165 170 175

Phe Ala Thr Met Ala Trp Ala Val Pro Lys Asp Asn Asn Lys Asn Ala

```
            180                 185                 190

Thr Asn Pro Leu Thr Val Glu Val Pro Tyr Ile Cys Thr Glu Gly Glu
        195                 200                 205

Asp Gln Ile Thr Val Trp Gly Phe His Ser Asp Asn Lys Thr Gln Met
    210                 215                 220

Lys Asn Leu Tyr Gly Asp Ser Asn Pro Gln Lys Phe Thr Ser Ser Ala
225                 230                 235                 240

Asn Gly Val Thr Thr His Tyr Val Ser Gln Ile Gly Ser Phe Pro Asp
                245                 250                 255

Gln Thr Glu Asp Gly Gly Leu Pro Gln Ser Gly Arg Ile Val Val Asp
            260                 265                 270

Tyr Met Met Gln Lys Pro Gly Lys Thr Gly Thr Ile Val Tyr Gln Arg
        275                 280                 285

Gly Val Leu Leu Pro Gln Lys Val Trp Cys Ala Ser Gly Arg Ser Lys
    290                 295                 300

Val Ile Lys Gly Ser Leu Pro Leu Ile Gly Glu Ala Asp Cys Leu His
305                 310                 315                 320

Glu Lys Tyr Gly Gly Leu Asn Lys Ser Lys Pro Tyr Tyr Thr Gly Glu
                325                 330                 335

His Ala Lys Ala Ile Gly Asn Cys Pro Ile Trp Val Lys Thr Pro Leu
            340                 345                 350

Lys Leu Ala Asn Gly Thr Lys Tyr Arg Pro Pro Gly Gly Gly Trp Glu
        355                 360                 365

Gly Met Ile Ala Gly Trp His Gly Tyr Thr Ser His Gly Ala His Gly
    370                 375                 380

Val Ala Val Ala Ala Asp Leu Lys Ser Thr Gln Glu Ala Ile Asn Lys
385                 390                 395                 400
```

Ile Thr Lys Asn Leu Asn Ser Leu Ser Glu Leu Glu Val Lys Asn Leu
405 410 415

Gln Arg Leu Ser Gly Ala Met Asp Glu Leu His Asn Glu Ile Leu Glu
420 425 430

Leu Asp Glu Lys Val Asp Asp Leu Arg Ala Asp Thr Ile Ser Ser Gln
435 440 445

Ile Glu Leu Ala Val Leu Leu Ser Asn Glu Gly Ile Ile Asn Ser Glu
450 455 460

Asp Glu His Leu Leu Ala Leu Glu Arg Lys Leu Lys Lys Met Leu Gly
465 470 475 480

Pro Ser Ala Val Glu Ile Gly Asn Gly Cys Phe Glu Thr Lys His Lys
485 490 495

Cys Asn Gln Thr Cys Leu Asp Arg Ile Ala Ala Gly Thr Phe Asn Ala
500 505 510

Gly Glu Phe Ser Leu Pro Thr Phe Asp Ser Leu Asn Ile Thr Ala Ala
515 520 525

Ser Leu Asn Asp Asp Gly Leu Asp Asn Tyr Gln Ile Leu Ala Ile Tyr
530 535 540

Ser Thr Val Ala Ser Ser Leu Val Leu Val Val Ser Leu Gly Ala Ile
545 550 555 560

Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
565 570 575

<210> 200
<211> 3486
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 2106

<400> 200

gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60

gaagaccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120

ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180

tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc tgccgacagt 240

ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc 300

acgtcttcaa agcaagtgga ttgatgtgat aacatggtgg agcacgacac acttgtctac 360

tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420

agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480

aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggt cccaaagatg gacccccacc cacgaggagc 600

atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660

tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720

taaggaagtt catttcattt ggagaggtat taaaatctta ataggttttg ataaaagcga 780

acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840

cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900

cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960

ggcgccatta ataacgtgt acttgtccta ttcttgtcgg tgtggtcttg ggaaaagaaa 1020

gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080

gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgcctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtattttc tttgaaacag agttttcccg 1200

tggttttcga acttggagaa agattgttaa gcttctgtat attctgccca aatttgtcgg 1260

gcccaatacc gcggagaaaa tggcgaaaaa cgttgcgatt ttcggcttat tgttttctct 1320

tcttgtgttg gttccttctc agatcttcgc ggatcgaatc tgcactggaa taacatcttc 1380

aaactcacct catgtggtca aaacagccac tcaaggggag gtcaatgtga ctggtgtgat 1440

accactaaca acaacaccaa caaaatctta ttttgcaaat ctcaaaggaa caaggaccag 1500

agggaaacta tgcccagact gtctcaactg cacagatctg gatgtggctt tgggcagacc 1560

aatgtgtgtg gggaccacac cttcggcgaa ggcttcaata ctccacgaag tcaaacctgt 1620

tacatccggg tgctttccta taatgcacga cagaacaaaa atcaggcaac tacccaatct 1680

tctcagagga tatgaaaata tcaggctatc aacccaaaac gtcatcgatg cggaaaaggc 1740

accaggagga ccctacagac ttggaacctc aggatcttgc cctaacgcta ccagtaagag 1800

cggatttttc gcaacaatgg cttgggctgt cccaaaggac aacaacaaaa atgcaacgaa 1860

cccactaaca gtagaagtac catacatttg tacagaaggg gaagaccaaa tcactgtttg 1920

ggggttccat tcagataaca aaacccaaat gaagaacctc tatggagact caaatcctca 1980

aaagttcacc tcatctgcta atggagtaac cacacactat gtttctcaga ttggcagctt 2040

cccagatcaa acagaagacg gaggactacc acaaagcggc aggattgttg ttgattacat 2100

gatgcaaaaa cctgggaaaa caggaacaat tgtctaccaa agaggtgttt tgttgcctca 2160

aaaggtgtgg tgcgcgagtg gcaggagcaa agtaataaaa gggtccttgc ctttaattgg 2220

tgaagcagat tgccttcatg aaaaatacgg tggattaaac aaaagcaagc cttactacac 2280

aggagaacat gcaaaagcca taggaaattg cccaatatgg gtgaaaacac ctttgaagct 2340

cgccaatgga accaaatata gacctcctgg tggaggatgg gaaggaatga ttgcaggctg 2400

gcacggatac acatctcacg gagcacatgg agtggcagtg gcggcggacc ttaagagtac 2460

gcaagaagct ataaacaaga taacaaaaaa tctcaattct ttgagtgagc tagaagtaaa 2520

gaatcttcaa agactaagtg gtgccatgga tgaactccac aacgaaatac tcgagctgga 2580

tgagaaagtg gatgatctca gagctgacac tataagctcg caaatagaac ttgcagtctt 2640

gctttccaac gaaggaataa taaacagtga agatgagcat ctattggcac ttgagagaaa 2700

actaaagaaa atgctgggtc cctctgctgt agagatagga aatggatgct tcgaaaccaa 2760

acacaagtgc aaccagacct gcttagacag gatagctgct ggcaccttta atgcaggaga 2820

attttctctc cccacttttg attcactgaa cattactgct gcatctttaa atgatgatgg 2880

attggataac taccagattt tggcgatcta ttcaactgtc gccagttcat tggtactggt 2940

agtctccctg ggggcaatca gtttctggat gtgctctaat gggtctctac agtgtagaat 3000

atgtatttaa aggcctattt tctttagttt gaatttactg ttattcggtg tgcatttcta 3060

tgtttggtga gcggttttct gtgctcagag tgtgtttatt ttatgtaatt taatttcttt 3120

gtgagctcct gtttagcagg tcgtcccttc agcaaggaca caaaaagatt ttaattttat 3180

taaaaaaaaa aaaaaaaaag accgggaatt cgatatcaag cttatcgacc tgcagatcgt 3240

tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt 3300

atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg 3360

ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta atacgcgata 3420

gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta 3480

ctagat 3486

<210> 201
<211> 5368
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Cátxét 2108

<400> 201
tgttgttgtg actccgaggg gttgcctcaa actctatctt ataaccggcg tggaggcatg 60

gaggcagggg tattttggtc attttaatag atagtggaaa atgacgtgga atttacttaa 120

agacgaagtc tttgcgacaa ggggggggccc acgccgaatt taatattacc ggcgtggccc 180

ccccttatcg cgagtgcttt agcacgagcg gtccagattt aaagtagaaa atttcccgcc 240

cactagggtt aaaggtgttc acactataaa agcatatacg atgtgatggt atttggtcga 300

caagcttgca tgccggtcaa catggtggag cacgacacac ttgtctactc caaaaatatc 360

aaagatacag tctcagaaga ccaaagggca attgagactt ttcaacaaag ggtaatatcc 420

ggaaacctcc tcggattcca ttgcccagct atctgtcact ttattgtgaa gatagtggaa 480

aaggaaggtg gctcctacaa atgccatcat tgcgataaag gaaaggccat cgttgaagat 540

gcctctgccg acagtggtcc caaagatgga cccccaccca cgaggagcat cgtggaaaaa 600

gaagacgttc caaccacgtc ttcaaagcaa gtggattgat gtgataacat ggtggagcac 660

gacacacttg tctactccaa aaatatcaaa gatacagtct cagaagacca aagggcaatt 720

gagacttttc aacaaagggt aatatccgga aacctcctcg gattccattg cccagctatc 780

tgtcacttta ttgtgaagat agtggaaaag gaaggtggct cctacaaatg ccatcattgc 840

gataaaggaa aggccatcgt tgaagatgcc tctgccgaca gtggtcccaa agatggaccc 900

ccacccacga ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc aaagcaagtg 960

gattgatgtg atatctccac tgacgtaagg gatgacgcac aatcccacta tccttcgcaa 1020

gacccttcct ctatataagg aagttcattt catttggaga ggtattaaaa tcttaatagg 1080

ttttgataaa agcgaacgtg gggaaacccg aaccaaacct tcttctaaac tctctctcat 1140

ctctcttaaa gcaaacttct ctcttgtctt tcttgcgtga gcgatcttca acgttgtcag 1200

atcgtgcttc ggcaccagta caacgttttc tttcactgaa gcgaaatcaa agatctcttt 1260

gtggacacgt agtgcggcgc cattaaataa cgtgtacttg tcctattctt gtcggtgtgg 1320

tcttgggaaa agaaagcttg ctggaggctg ctgttcagcc ccatacatta cttgttacga 1380

ttctgctgac tttcggcggg tgcaatatct ctacttctgc ttgacgaggt attgttgcct 1440

gtacttcttt cttcttcttc ttgctgattg gttctataag aaatctagta ttttctttga 1500

aacagagttt tcccgtggtt ttcgaacttg gagaaagatt gttaagcttc tgtatattct 1560

gcccaaattt gtcgggccca ataccgcgga gaaaatggcg aaaaacgttg cgattttcgg 1620

cttattgttt tctcttcttg tgttggttcc ttctcagatc ttcgcggatc gaatctgcac 1680

tggaataaca tcttcaaact cacctcatgt ggtcaaaaca gccactcaag gggaggtcaa 1740

tgtgactggt gtgataccac taacaacaac accaacaaaa tcttattttg caaatctcaa 1800

aggaacaagg accagaggga aactatgccc agactgtctc aactgcacag atctggatgt 1860

ggctttgggc agaccaatgt gtgtggggac cacaccttcg gcgaaggctt caatactcca 1920

cgaagtcaaa cctgttacat ccgggtgctt tcctataatg cacgacagaa caaaaatcag 1980

gcaactaccc aatcttctca gaggatatga aaatatcagg ctatcaaccc aaaacgtcat 2040

cgatgcggaa aaggcaccag gaggacccta cagacttgga acctcaggat cttgccctaa 2100

cgctaccagt aagagcggat ttttcgcaac aatggcttgg gctgtcccaa aggacaacaa 2160

caaaaatgca acgaacccac taacagtaga agtaccatac atttgtacag aaggggaaga 2220

ccaaatcact gtttgggggt tccattcaga taacaaaacc caaatgaaga acctctatgg 2280

agactcaaat cctcaaaagt tcacctcatc tgctaatgga gtaaccacac actatgtttc 2340

tcagattggc agcttcccag atcaaacaga agacggagga ctaccacaaa gcggcaggat 2400

tgttgttgat tacatgatgc aaaaacctgg gaaaacagga acaattgtct accaaagagg 2460

tgttttgttg cctcaaaagg tgtggtgcgc gagtggcagg agcaaagtaa taaaagggtc 2520

cttgccttta attggtgaag cagattgcct tcatgaaaaa tacggtggat taaacaaaag 2580

caagccttac tacacaggag aacatgcaaa agccatagga aattgcccaa tatgggtgaa 2640

aacacctttg aagctcgcca atggaaccaa atatagacct cctggtggag gatgggaagg 2700

aatgattgca ggctggcacg gatacacatc tcacggagca catggagtgg cagtggcggc 2760

ggaccttaag agtacgcaag aagctataaa caagataaca aaaaatctca attctttgag 2820

tgagctagaa gtaaagaatc ttcaaagact aagtggtgcc atggatgaac tccacaacga 2880

aatactcgag ctggatgaga aagtggatga tctcagagct gacactataa gctcgcaaat 2940

agaacttgca gtcttgcttt ccaacgaagg aataataaac agtgaagatg agcatctatt 3000

ggcacttgag agaaaactaa agaaaatgct gggtccctct gctgtagaga taggaaatgg 3060

atgcttcgaa accaaacaca agtgcaacca gacctgctta gacaggatag ctgctggcac 3120

ctttaatgca ggagaatttt ctctccccac ttttgattca ctgaacatta ctgctgcatc 3180

tttaaatgat gatggattgg ataactacca gattttggcg atctattcaa ctgtcgccag 3240

ttcattggta ctggtagtct ccctgggggc aatcagtttc tggatgtgct ctaatgggtc 3300

tctacagtgt agaatatgta tttaaaggcc tattttcttt agtttgaatt tactgttatt 3360

cggtgtgcat ttctatgttt ggtgagcggt tttctgtgct cagagtgtgt ttattttatg 3420

taatttaatt tctttgtgag ctcctgttta gcaggtcgtc ccttcagcaa ggacacaaaa 3480

agattttaat tttattaaaa aaaaaaaaaa aaaagaccgg gaattcgata tcaagcttat 3540

cgacctgcag atcgttcaaa catttggcaa taaagtttct taagattgaa tcctgttgcc 3600

ggtcttgcga tgattatcat ataatttctg ttgaattacg ttaagcatgt aataattaac 3660

atgtaatgca tgacgttatt tatgagatgg gtttttatga ttagagtccc gcaattatac 3720

atttaatacg cgatagaaaa caaaatatag cgcgcaaact aggataaatt atcgcgcgcg 3780

gtgtcatcta tgttactaga tctctagagt ctcaagcttg gcgcgccata aaatgattat 3840

tttatgaata tatttcattg tgcaagtaga tagaaattac atatgttaca taacacacga 3900

aataaacaaa aaaagacaat ccaaaaacaa acaccccaaa aaaaataatc actttagata 3960

aactcgtatg aggagaggca cgttcagtga ctcgacgatt cccgagcaaa aaaagtctcc 4020

ccgtcacaca tatagtgggt gacgcaatta tctttaaagt aatccttctg ttgacttgtc 4080

attgataaca tccagtcttc gtcaggattg caaagaatta tagaagggat cccacctttt 4140

attttcttct tttttccata tttagggttg acagtgaaat cagactggca acctattaat 4200

tgcttccaca atgggacgaa cttgaagggg atgtcgtcga tgatattata ggtggcgtgt 4260

tcatcgtagt tggtgaaatc gatggtaccg ttccaatagt tgtgtcgtcc gagacttcta 4320

gcccaggtgg tctttccggt acgagttggt ccgcagatgt agaggctggg gtgtcggatt 4380

ccattccttc cattgtcctg gttaaatcgg ccatccattc aaggtcagat tgagcttgtt 4440

ggtatgagac aggatgtatg taagtataag cgtctatgct tacatggtat agatgggttt 4500

ccctccagga gtgtagatct tcgtggcagc gaagatctga ttctgtgaag ggcgacacat 4560

acggttcagg ttgtggaggg aataatttgt tggctgaata ttccagccat tgaagttttg 4620

ttgcccattc atgagggaat tcttccttga tcatgtcaag atattcctcc ttagacgttg 4680

cagtctggat aatagttctc catcgtgcgt cagatttgcg aggagagacc ttatgatctc 4740

ggaaatctcc tctggtttta atatctccgt cctttgatat gtaatcaagg acttgtttag 4800

agtttctagc tggctggata ttagggtgat ttccttcaaa atcgaaaaaa gaaggatccc 4860

taatacaagg ttttttatca agctggagaa gagcatgata gtgggtagtg ccatcttgat 4920

gaagctcaga agcaacacca aggaagaaaa taagaaaagg tgtgagtttc tcccagagaa 4980

actggaataa atcatctctt tgagatgagc acttgggata ggtaaggaaa acatatttag 5040

attggagtct gaagttctta ctagcagaag gcatgttgtt gtgactccga ggggttgcct 5100

caaactctat cttataaccg gcgtggaggc atggaggcag gggtattttg gtcattttaa 5160

tagatagtgg aaaatgacgt ggaatttact taaagacgaa gtctttgcga caaggggggg 5220

cccacgccga atttaatatt accggcgtgg ccccccctta tcgcgagtgc tttagcacga 5280

gcggtccaga tttaaagtag aaaatttccc gcccactagg gttaaaggtg ttcacactat 5340

aaaagcatat acgatgtgat ggtatttg 5368

Fig.1: A-2X35S/CPMV-HT/H5 Indonesia/NOS (Cấu trúc số 489)

Fig.1A

IF-H5A-I-05.s1+3c SEQ ID NO: 2

AAATTTGTCGGGCCCATGGAGAAAATAGTGCTTCTTCTTGC

Fig.1B

IF-H5dTm.r SEQ ID NO: 3

ACTAAAGAAAATAGGCCTTTAAATGCAAATTCTGCATTGTAACGATCCAT

Fig.1C

Sự biểu diễn theo sơ đồ của cấu trúc 1191. Vị trí enzym giới hạn SacII và StuI được sử dụng để tuyến tính hoá plasmit được chú giải đối với sự biểu hiện.

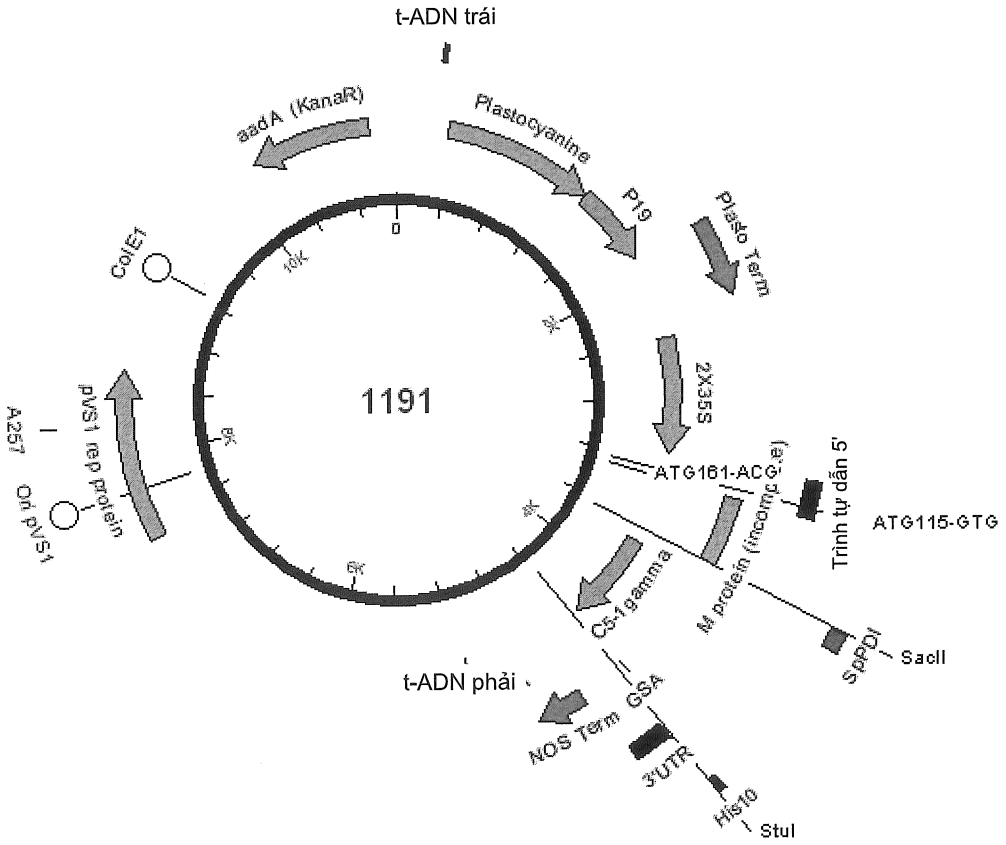


Fig.1D
Cấu trúc 1191 từ bên trái sang bên phải đường biên t-ADN (gạch dưới).
2X35S/CPMV-HT/NOS với Plastoxyanin-P19-Plastoxyanin làm câm cát xét biểu hiện chất ức chế (SEQ ID NO:4)

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACTTAATAACACATTGCGGACG
TTTTTAATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTC
AAAAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTA
CTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAG
TGATATTTTGACAACAATTTTGTTGCAACATTTGAGAAAATTTTGTTGTTCTCTCTTTTCATTG
GTCAAAAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGA
GAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGC
TACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTA
GAGAATTTTTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTTAAAATTAAAAGTTGAGTCA
TTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAA
TTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCCATAGAGTC
AGTTAACTCATTTTTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAA
AAAAAAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTA
GGAGGATAACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCAC
GCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAA
ACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATT
CCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGA
ACGAGCTATACAAGGAAACGACGCTAGGGAACAAGCTAACAGTGAACGTTGGGATGGAGGATCA
GGAGGTACCACTTCTCCCTTCAAACTTCCTGACGAAAGTCCGAGTTGGACTGAGTGGCGGCTAC
ATAACGATGAGACGAATTCGAATCAAGATAATCCCCTTGGTTTCAAGGAAAGCTGGGGTTTCGG
GAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCACTGCACAGAGTCCTT
GGATCTTGGACGGGAGATTCGGTTAACTATGCAGCATCTCGATTTTTCGGTTTCGACCAGATCG
GATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACTCT
TCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAA
GTGGAAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGT
AAGTTAAAATGCTTCTTCGTCTCCTATTTATAATATGGTTTGTTATTGTTAATTTTGTTCTTGT
AGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTGGTGTAATG
TAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCCATAACTAACTAGACATGAAG
ACCTGCCGCGTACAATTGTCTTATATTTGAACAACTAAAATTGAACATCTTTTGCCACAACTTT
ATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGT
TATCGAAATTCATTAACAATCAACTTAACGTTATTAACTACTAATTTTATATCATCCCCTTTGA
TAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTGTCGTTTCCCGCCTTC
AGTTTGCAAGCTGCTCTAGCCGTGTAGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGGAAT
TACTAGCGCGTGTCGACAAGCTTGCATGCCGGTCAACATGGTGGAGCACGACACACTTGTCTAC
TCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGG
TAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGT
GGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGAT
GCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAG
ACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACT
TGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAA
CAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGA
AGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGT

TGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAA
AAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAA
GGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCA
TTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAAACCCGAACCAA
ACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGA
GCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGA
AATCAAAGATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTC
TTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTCAGCCCCATACATTACT
TGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTT
GCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGA
AACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCC
AAATTTGTCGGGCCCGCGGATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTT
GTGTTGGTTCCTTCTCAGATCTTCGCCTGCAGGCTCCTCAGCCAAAACGACACCCCCATCTGTC
TATCCACTGGCCCCTGGATCTGCTGCCCAAACTAACTCCATGGTGACCCTGGGATGCCTGGTCA
AGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACTCTGGATCCCTGTCCAGCGGTGTGCA
CACCTTCCCAGCTGTCCTGCAGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGTCCCCTCC
AGCACCTGGCCCAGCGAGACCGTCACCTGCAACGTTGCCCACCCGGCCAGCAGCACCAAGGTGG
ACAAGAAAATTGTGCCCAGGGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATC
ATCTGTCTTCATCTTCCCCCCAAAGCCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTC
ACGTGTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTAGATG
ATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGCTC
AGTCAGTGAACTTCCCATCATGCACCAGGACTGGCTCAATGGCAAGGAGCGATCGCTCACCATC
ACCATCACCATCACCATCACCATTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGG
TGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTA
ATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATT
TTATTAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGT
TCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCA
TATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTAT
GAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATA
TAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTC
TCAAGCTTGGCGCGCCCACGTGACTAGTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGA
AAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAAT
AGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCTAGA
GCAGCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGACAG
GATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTTTA

Fig.1E
Cát xét biểu hiện số 489 từ đoạn khởi đầu 2X35S đến kết thúc NOS. H5 từ cúm A/Indonesia/5/2005 (H5N1) được gạch dưới. SEQ NO:5

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAG
ACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCAT
CATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGAC
CCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGA
TTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTC
TCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGAT
TCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAA
ATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAA
GATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGC
AAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCA
AGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTT
TGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTT
AAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCG
GCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGC
GGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTT
GCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGC
AATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGA
TTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGG
AGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGGAGAAAATAGTGC
TTCTTCTTGCAATAGTCAGTCTTGTTAAAAGTGATCAGATTTGCATTGGTTACCATGCAAACAA
TTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGTTACTGTTACACATGCCCAAGACATA
CTGGAAAAGACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTTAAGAG
ATTGTAGTGTAGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCGGA
ATGGTCTTACATAGTGGAGAAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGAGTTTCAAC
GACTATGAAGAACTGAAACACCTATTGAGCAGAATAAACCATTTTGAGAAAATTCAAATCATCC
CCAAAAGTTCTTGGTCCGATCATGAAGCCTCATCAGGAGTTAGCTCAGCATGTCCATACCTGGG
AAGTCCCTCCTTTTTTAGAAATGTGGTATGGCTTATCAAAAAGAACAGTACATACCCAACAATA
AAGAAAAGCTACAATAATACCAACCAAGAGGATCTTTTGGTACTGTGGGGAATTCACCATCCTA
ATGATGCGGCAGAGCAGACAAGGCTATATCAAAACCCAACCACCTATATTTCCATTGGGACATC
AACACTAAACCAGAGATTGGTACCAAAAATAGCTACTAGATCCAAAGTAAACGGGCAAAGTGGA
AGGATGGAGTTCTTCTGGACAATTTTAAAACCTAATGATGCAATCAACTTCGAGAGTAATGGAA
ATTTCATTGCTCCAGAATATGCATACAAAATTGTCAAGAAAGGGGACTCAGCAATTATGAAAAG
TGAATTGGAATATGGTAACTGCAACACCAAGTGTCAAACTCCAATGGGGGCGATAAACTCTAGT
ATGCCATTCCACAACATACACCCTCTCACCATCGGGGAATGCCCCAAATATGTGAAATCAAACA
GATTAGTCCTTGCAACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAAGAAAAAAGAGAGG
ACTATTTGGAGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTAT
GGGTACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTCAAAAGG
CAATAGATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGC
CGTTGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGAC
GGGTTTCTAGATGTCTGGACTTATAATGCCGAACTTCTGGTTCTCATGGAAAATGAGAGAACTC
TAGACTTTCATGACTCAAATGTTAAGAACCTCTACGACAAGGTCCGACTACAGCTTAGGGATAA

TGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATCACAAATGTGATAATGAATGTATGGAA
AGTATAAGAAACGGAACGTACAACTATCCGCAGTATTCAGAAGAAGCAAGATTAAAAAGAGAGG
AAATAAGTGGGGTAAAATTGGAATCAATAGGAACTTACCAAATACTGTCAATTTATTCAACAGT
GGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGA
TCGTTACAATGCAGAATTTGCATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCG
GTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTT
AATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAT
TTTATTAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCG
TTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATC
ATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTA
TGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAAT
ATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.1F
Trình tự axit amin của H5 từ cúm A/Indonesia/5/2005 (H5N1) SEQ ID NO:6

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCDLDGVK
PLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGSFNDYEELKHLLSRINHFE
KIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIKKSYNNTNQEDLLVLW
GIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVNGQSGRMEFFWTILKPNDAIN
FESNGNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTKCQTPMGAINSSMPFHNIHPLTIGECPK
YVKSNRLVLATGLRNSPQRESRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADK
ESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLNKKMEDGFLDVWTYNAELLVLM
ENERTLDFHDSNVKNLYDKVRLQLRDNAKELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEA
RLKREEISGVKLESIGTYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

Fig.1G
Trình tự nucleotit mã hoá H5 từ cúm A/Indonesia/5/2005 (H5N1) SEQ ID NO:42

ATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCAGTCTTGTTAAAAGTGATCAGATTTGCATTG
GTTACCATGCAAACAATTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGTTACTGTTAC
ACATGCCCAAGACATACTGGAAAAGACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAG
CCTCTAATTTTAAGAGATTGTAGTGTAGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAAT
TCATCAATGTACCGGAATGGTCTTACATAGTGGAGAAGGCCAATCCAACCAATGACCTCTGTTA
CCCAGGGAGTTTCAACGACTATGAAGAACTGAAACACCTATTGAGCAGAATAAACCATTTTGAG
AAAATTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCTCATCAGGAGTTAGCTCAG
CATGTCCATACCTGGGAAGTCCCTCCTTTTTTAGAAATGTGGTATGGCTTATCAAAAAGAACAG
TACATACCCAACAATAAAGAAAAGCTACAATAATACCAACCAAGAGGATCTTTTGGTACTGTGG
GGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGCTATATCAAAACCCAACCACCTATA
TTTCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAAAATAGCTACTAGATCCAAAGT
AAACGGGCAAAGTGGAAGGATGGAGTTCTTCTGGACAATTTTAAAACCTAATGATGCAATCAAC
TTCGAGAGTAATGGAAATTTCATTGCTCCAGAATATGCATACAAAATTGTCAAGAAAGGGGACT
CAGCAATTATGAAAAGTGAATTGGAATATGGTAACTGCAACACCAAGTGTCAAACTCCAATGGG
GGCGATAAACTCTAGTATGCCATTCCACAACATACACCCTCTCACCATCGGGGAATGCCCCAAA
TATGTGAAATCAAACAGATTAGTCCTTGCAACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCA
GAAGAAAAAAGAGAGGACTATTTGGAGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAAT
GGTAGATGGTTGGTATGGGTACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAA
GAATCCACTCAAAAGGCAATAGATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGA
ACACTCAGTTTGAGGCCGTTGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGAGAATTTAAA
CAAGAAGATGGAAGACGGGTTTCTAGATGTCTGGACTTATAATGCCGAACTTCTGGTTCTCATG
GAAAATGAGAGAACTCTAGACTTTCATGACTCAAATGTTAAGAACCTCTACGACAAGGTCCGAC
TACAGCTTAGGGATAATGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATCACAAATGTGA
TAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAACTATCCGCAGTATTCAGAAGAAGCA
AGATTAAAAAGAGAGGAAATAAGTGGGGTAAAATTGGAATCAATAGGAACTTACCAAATACTGT
CAATTTATTCAACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATG
GATGTGCTCCAATGGATCGTTACAATGCAGAATTTGCATTTAA

Fig 2: B-2X35S/CPMV HT/M2 Caledonia/NOS mới (Cấu trúc số 1261)

Fig.2A

IF-S1-M1+M2ANC.c (SEQ ID NO:7)

AAATTTGTCGGGCCCATGAGTCTTCTAACCGAGGTCGAAACG

Fig.2B

IF-S1-4-M2ANC.r (SEQ ID NO: 8)

ACTAAAGAAAATAGGCCTTTACTCCAGCTCTATGCTGACAAAA

Fig.2C
Gen M2 tổng hợp (tương ứng với nt 1-26 kết hợp với 715-982 từ số truy cập Genbank DQ508860) (SEQ ID NO:9)

ATGAGTCTTCTAACCGAGGTCGAAACGCCTATCAGAAACGAATGGGGGTGCAGATGCAACGATT
CAAGTGATCCTCTTGTTGTTGCCGCAAGTATAATTGGGATTGTGCACCTGATATTGTGGATTAT
TGATCGCCTTTTTTCCAAAAGCATTTATCGTATCTTTAAACACGGTTTAAAAAGAGGGCCTTCT
ACGGAAGGAGTACCAGAGTCTATGAGGGAAGAATATCGAGAGGAACAGCAGAATGCTGTGGATG
CTGACGATGGTCATTTTGTCAGCATAGAGCTGGAGTAA

Fig.2D
Cát xét biểu hiện số 1261 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. M2 từ cúm A/Caledonia/20/1999 mới (H1N1) được gạch dưới. (SEQ ID NO:10)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAG
ACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCAT
CATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGAC
CCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGA
TTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTC
TCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGAT
TCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAA
ATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAA
GATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGC
AAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCA

AGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTT
TGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTT
AAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCG
GCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGC
GGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTT
GCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGC
AATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGA
TTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGG
AGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGAGTCTTCTAACCG
AGGTCGAAACGCCTATCAGAAACGAATGGGGGTGCAGATGCAACGATTCAAGTGATCCTCTTGT
TGTTGCCGCAAGTATAATTGGGATTGTGCACCTGATATTGTGGATTATTGATCGCCTTTTTTCC
AAAAGCATTTATCGTATCTTTAAACACGGTTTAAAAAGAGGGCCTTCTACGGAAGGAGTACCAG
AGTCTATGAGGGAAGAATATCGAGAGGAACAGCAGAATGCTGTGGATGCTGACGATGGTCATTT
TGTCAGCATAGAGCTGGAGTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTG
CATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTT
CTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTAT
TAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAA
ACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATA
ATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGA
TGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGC
GCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.2E
Trình tự axit amin của M2 từ cúm A/Caledonia/20/1999 mới (H1N1) (SEQ ID NO:11)

MSLLTEVETPIRNEWGCRCNDSSDPLVVAASIIGIVHLILWIIDRLFSKSIYRIFKHGLKRGPS
TEGVPESMREEYREEQQNAVDADDGHFVSIELE

Fig.3: C-2X35S/CPMV-HT/M2 Puerto Rico/NOS (Cấu trúc số 859)

Fig.3A
Gen M2 tổng hợp (tương ứng với ni 26-51 kết hợp với 740-1007 từ số truy cập Genbank EF467824) (SEQ ID NO:12)

ATGAGTCTTCTAACCGAGGTCGAAACGCCTATCAGAAACGAATGGGGGTGCAGATGCAACGGTT
CAAGTGATCCTCTCACTATTGCCGCAAATATCATTGGGATCTTGCACTTGACATTGTGGATTCT
TGATCGTCTTTTTTTCAAATGCATTTACCGTCGCTTTAAATACGGACTGAAAGGAGGGCCTTCT
ACGGAAGGAGTGCCAAAGTCTATGAGGGAAGAATATCGAAAGGAACAGCAGAGTGCTGTGGATG
CTGACGATGGTCATTTTGTCAGCATAGAGCTGGAGTAA

Fig.3B

Cát xét biểu hiện số 859 từ đoạn khởi đầu 2X35S đến kết thúc NOS. M2 từ cúm A/Puerto Rico/8/1934 (H1N1) được gạch dưới. (SEQ ID NO:13)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAG
ACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCAT
CATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGAC
CCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGA
TTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTC
TCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGAT
TCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAA
ATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAA
GATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGC
AAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCA
AGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTT
TGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTT
AAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCG
GCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGC
GGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTT
GCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGC
AATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGA
TTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGG
AGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGAGTCTTCTAACCG
AGGTCGAAACGCCTATCAGAAACGAATGGGGGTGCAGATGCAACGGTTCAAGTGATCCTCTCAC
TATTGCCGCAAATATCATTGGGATCTTGCACTTGACATTGTGGATTCTTGATCGTCTTTTTTTC
AAATGCATTTACCGTCGCTTTAAATACGGACTGAAAGGAGGGCCTTCTACGGAAGGAGTGCCAA
AGTCTATGAGGGAAGAATATCGAAAGGAACAGCAGAGTGCTGTGGATGCTGACGATGGTCATTT
TGTCAGCATAGAGCTGGAGTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTG
CATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTT

CTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTAT
TAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAA
ACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATA
ATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGA
TGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGC
GCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.3C

Trình tự axit amin của M2 từ cúm A/Puerto Rico/8/1934 (H1N1) (SEQ ID NO:14)

MSLLTEVETPIRNEWGCRCNGSSDPLTIAANIIGILHLTLWILDRLFFKCIYRRFKYGLKGGPS
TEGVPKSMREEYRKEQQSAVDADDGHFVSIELE

Fig. 4: G-2X35S/CPMV-HT/PDISP/HA B Brisbane/NOS vào trong hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1008)

Fig.4A
Sự biểu diễn theo sơ đồ của cấu trúc 1194. Vị trí enzym giới hạn SacII và StuI được sử dụng để tuyến tính hoá plasmit được chú giải đối với sự biểu diễn.

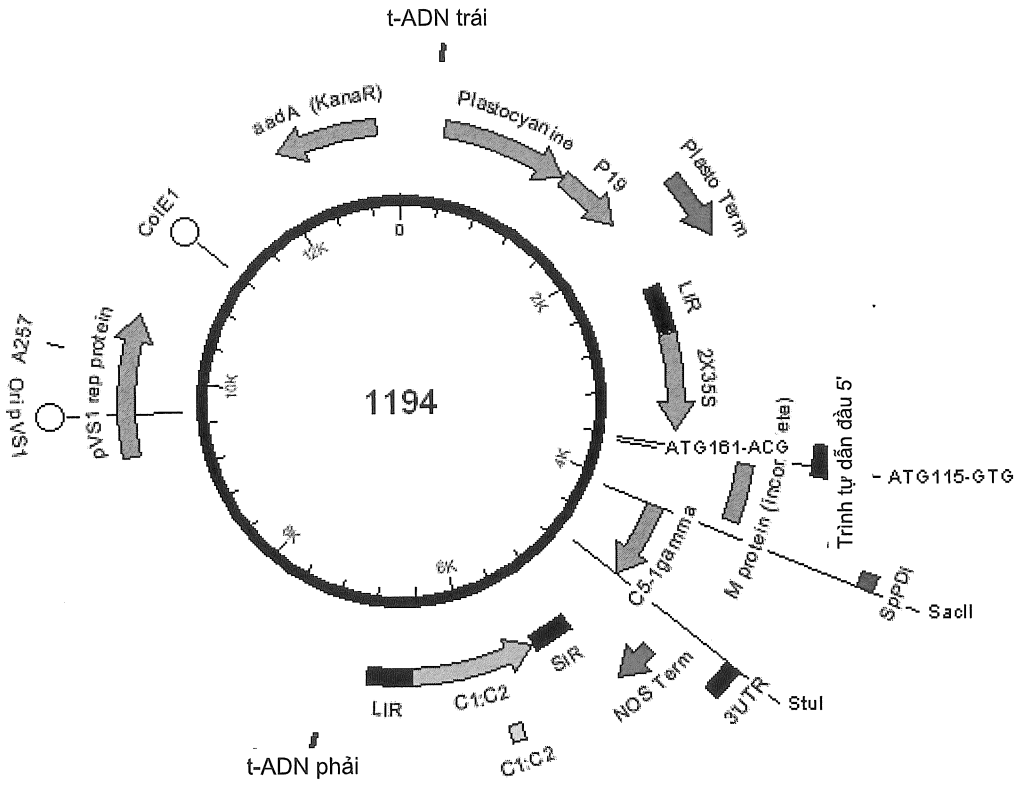


Fig.4B

Cấu trúc 1194 từ bên trái sang bên phải đường biên t-ADN (gạch dưới).
2X35S/CPMV-HT/PDISP/NOS vào trong hệ khuếch đại biểu hiện BeYDV+Replicaza với Plastoxyanin-P19-Plastoxyanin làm câm cát xét biểu hiện chất ức chế (SEQ ID NO:31)

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACTTAATAACACATTGCGGACG
TTTTTAATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTC
AAAAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTA
CTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAG
TGATATTTTGACAACAATTTTGTTGCAACATTTGAGAAAATTTTGTTGTTCTCTCTTTTCATTG
GTCAAAAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGA
GAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGC

TACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTA
GAGAATTTTTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTTAAAATTAAAAGTTGAGTCA
TTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAA
TTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCCATAGAGTC
AGTTAACTCATTTTTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAA
AAAAAAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTA
GGAGGATAACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCAC
GCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAA
ACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATT
CCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGA
ACGAGCTATACAAGGAAACGACGCTAGGGAACAAGCTAACAGTGAACGTTGGGATGGAGGATCA
GGAGGTACCACTTCTCCCTTCAAACTTCCTGACGAAAGTCCGAGTTGGACTGAGTGGCGGCTAC
ATAACGATGAGACGAATTCGAATCAAGATAATCCCCTTGGTTTCAAGGAAAGCTGGGGTTTCGG
GAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCACTGCACAGAGTCCTT
GGATCTTGGACGGGAGATTCGGTTAACTATGCAGCATCTCGATTTTTCGGTTTCGACCAGATCG
GATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACTCT
TCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAA
GTGGAAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGT
AAGTTAAAATGCTTCTTCGTCTCCTATTTATAATATGGTTTGTTATTGTTAATTTTGTTCTTGT
AGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTGGTGTAATG
TAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCCATAACTAACTAGACATGAAG
ACCTGCCGCGTACAATTGTCTTATATTTGAACAACTAAAATTGAACATCTTTTGCCACAACTTT
ATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGT
TATCGAAATTCATTAACAATCAACTTAACGTTATTAACTACTAATTTTATATCATCCCCTTTGA
TAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTGTCGTTTCCCGCCTTC
AGTTTGCAAGCTGCTCTAGCCGTGTAGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGGAAT
TACTAGCGCGTGTCGACACGCGTGGCGCGCCCTAGCAGAAGGCATGTTGTTGTGACTCCGAGGG
GTTGCCTCAAACTCTATCTTATAACCGGCGTGGAGGCATGGAGGCAAGGGCATTTTGGTAATTT
AAGTAGTTAGTGGAAAATGACGTCATTTACTTAAAGACGAAGTCTTGCGACAAGGGGGGCCCAC
GCCGAATTTTAATATTACCGGCGTGGCCCCACCTTATCGCGAGTGCTTTAGCACGAGCGGTCCA
GATTTAAAGTAGAAAAGTTCCCGCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATAC
GATGTGATGGTATTTGATAAAGCGTATATTGTATCAGGTATTTCCGTCGGATACGAATTATTCG
TACAAGCTTCTTAAGCCGGTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCA
AAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAA
CCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGT
GGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACA
GTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCAC
GTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAA
AATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATAT
CCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAA
GGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCT
GCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTC
CAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACA
ATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTAT
TAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAA
CTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACG
TTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTC

TTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGT
CTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTG
CTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTT
TCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCC
CGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGC
CCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCA
GATCTTCGCCGCGGCTCCTCAGCCAAAACGACACCCCCATCTGTCTATCCACTGGCCCCTGGAT
CTGCTGCCCAAACTAACTCCATGGTGACCCTGGGATGCCTGGTCAAGGGCTATTTCCCTGAGCC
AGTGACAGTGACCTGGAACTCTGGATCCCTGTCCAGCGGTGTGCACACCTTCCCAGCTGTCCTG
CAGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGTCCCCTCCAGCACCTGGCCCAGCGAGA
CCGTCACCTGCAACGTTGCCCACCCGGCCAGCAGCACCAAGGTGGACAAGAAAATTGTGCCCAG
GGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCATCTTCCCC
CCAAAGCCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACA
TCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTAGATGATGTGGAGGTGCACACAGC
TCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGCTCAGTCAGTGAACTTCCCATC
ATGCACCAGGACTGGCTCAATGGCAAGGAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATT
CGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAAT
TTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTA
ATTTTATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGAT
CGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTA
TCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATT
TATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAA
ATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGA
GTCTCAAGCTTGGCGCGGGGTACCGAGCTCGAATTCCGAGTGTACTTCAAGTCAGTTGGAAATC
AATAAAATGATTATTTTATGAATATATTTCATTGTGCAAGTAGATAGAAATTACATATGTTACA
TAACACACGAAATAAACAAAAAAACACAATCCAAAACAAACACCCCAAACAAAATAACACTATA
TATATCCTCGTATGAGGAGAGGCACGTTCAGTGACTCGACGATTCCCGAGCAAAAAAAGTCTCC
CCGTCACACATATAGTGGGTGACGCAATTATCTTCAAAGTAATCCTTCTGTTGACTTGTCATTG
ATAACATCCAGTCTTCGTCAGGATTGCAAAGAATTATAGAAGGGATCCCACCTTTTATTTTCTT
CTTTTTTCCATATTTAGGGTTGACAGTGAAATCAGACTGGCAACCTATTAATTGCTTCCACAAT
GGGACGAACTTGAAGGGGATGTCGTCGATGATATTATAGGTGGCGTGTTCATCGTAGTTGGTGA
AGTCGATGGTCCCGTTCCAGTAGTTGTGTCGCCCGAGACTTCTAGCCCAGGTGGTCTTTCCGGT
ACGAGTTGGTCCGCAGATGTAGAGGCTGGGGTGTCTGACCCCAGTCCTTCCCTCATCCTGGTTA
GATCGGCCATCCACTCAAGGTCAGATTGTGCTTGATCGTAGGAGACAGGATGTATGAAAGTGTA
GGCATCGATGCTTACATGATATAGGTGCGTCTCTCTCCAGTTGTGCAGATCTTCGTGGCAGCGG
AGATCTGATTCTGTGAAGGGCGACACGTACTGCTCAGGTTGTGGAGGAAATAATTTGTTGGCTG
AATATTCCAGCCATTGAAGCTTTGTTGCCCATTCATGAGGGAATTCTTCTTTGATCATGTCAAG
ATACTCCTCCTTAGACGTTGCAGTCTGGATAATAGTTCGCCATCGTGCGTCAGATTTGCGAGGA
GAGACCTTATGATCTCGGAAATCTCCTCTGGTTTTAATATCTCCGTCCTTTGATATGTAATCAA
GGACTTGTTTAGAGTTTCTAGCTGGCTGGATATTAGGGTGATTTCCTTCAAAATCGAAAAAAGA
AGGATCCCTAATACAAGGTTTTTTATCAAGCTGGATAAGAGCATGATAGTGGGTAGTGCCATCT
TGATGAAGCTCAGAAGCAACACCAAGGAAGAAAATAAGAAAAGGTGTGAGTTTCTCCCAGAGAA
ACTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGGTAGGTAAGGAAAACATATTTAGATTG
GAGTCTGAAGTTCTTGCTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACTCT
ATCTTATAACCGGCGTGGAGGCATGGAGGCAAGGGCATTTTGGTAATTTAAGTAGTTAGTGGAA
AATGACGTCATTTACTTAAAGACGAAGTCTTGCGACAAGGGGGGCCCACGCCGAATTTTAATAT
TACCGGCGTGGCCCCACCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAA

AGTTCCCGCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGATGGTATTT
GATGGAGCGTATATTGTATCAGGTATTTCCGTCGGATACGAATTATTCGTACGGCCGGCCACTA
GTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATC
GCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCC
TTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGATTGTC
GTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAG
AAAAGAGCGTTTA

Fig.4C

Cát xét biểu hiện số 1008 từ BeYDV LIR bên trái sang BeYDV LIR bên phải. PDISP/HA từ cúm B/Brisbane/60/2008 được gạch dưới (SEQ ID NO:32)

CTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACTCTATCTTATAACCGGCGT
GGAGGCATGGAGGCAAGGGCATTTTGGTAATTTAAGTAGTTAGTGGAAAATGACGTCATTTACT
TAAAGACGAAGTCTTGCGACAAGGGGGGCCCACGCCGAATTTTAATATTACCGGCGTGGCCCCA
CCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAAGTTCCCGCCCACTAG
GGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGATGGTATTTGATAAAGCGTATATTG
TATCAGGTATTTCCGTCGGATACGAATTATTCGTACAAGCTTCTTAAGCCGGTCAACATGGTGG
AGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAAT
TGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGT
CACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAG
GAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAG
GAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAAC
ATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAA
GGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGC
TATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGC
GATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCAC
CCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATG
TGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCT
ATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAA
CGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCT
CTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAAC
GTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAAT
AACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCT
GTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTT
CTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAG
AAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTA
AGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTT
ATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCGAATCTGCACTGGAATA
ACATCGTCAAACTCACCACATGTCGTCAAAACTGCTACTCAAGGGGAGGTCAATGTGACTGGTG
TAATACCACTGACAACAACACCCACCAAATCTCATTTTGCAAATCTCAAAGGAACAGAAACCAG
GGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATCTGGACGTAGCCTTGGGCAGACCAAAA
TGCACGGGGAAAATACCCTCGGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTACATCTG
GGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCCTAACCTTCTCCGAGGATA
CGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAAAATGCACCAGGAGGACCCTAC

AAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTTCGCAACAATGG
CTTGGGCCGTCCCAAAAAACGACAAAAACAAAACAGCAACAAATCCATTAACAATAGAAGTACC
ATACATTTGTACAGAAGGAGAAGACCAAATTACCGTTTGGGGGTTCCACTCTGACAACGAGACC
CAAATGGCAAAGCTCTATGGGGACTCAAAGCCCCAGAAGTTCACCTCATCTGCCAACGGAGTGA
CCACACATTACGTTTCACAGATTGGTGGCTTCCCAAATCAAACAGAAGACGGAGGACTACCACA
AAGTGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGGGAAAACAGGAACAATTACCTAT
CAAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTGCGCAAGTGGCAGGAGCAAGGTAATAAAAG
GATCCTTGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTGGATTAAACAAAAG
CAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAATATGGGTGAAAACA
CCCTTGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGCAAAACTATTAAAGGAAAGGGGTT
TCTTCGGAGCTATTGCTGGTTTCTTAGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGG
ATACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCC
ATAAACAAGATAACAAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAGAC
TAAGCGGTGCCATGGATGAACTCCACAACGAAATACTAGAACTAGATGAGAAAGTGGATGATCT
CAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAGTCCTGCTTTCCAATGAAGGAATAATA
AACAGTGAAGATGAACATCTCTTGGCGCTTGAAAGAAAGCTGAAGAAAATGCTGGGCCCCTCTG
CTGTAGAGATAGGGAATGGATGCTTTGAAACCAAACACAAGTGCAACCAGACCTGTCTCGACAG
AATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTCTCCCCACCTTTGATTCACTGAATATT
ACTGCTGCATCTTTAAATGACGATGGATTGGATAATCATACTATACTGCTTTACTACTCAACTG
CTGCCTCCAGTTTGGCTGTAACACTGATGATAGCTATCTTTGTTGTTTATATGGTCTCCAGAGA
CAATGTTTCTTGCTCCATCTGTCTATAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTC
GGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATT
TAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAA
TTTTATTAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATC
GTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTAT
CATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTT
ATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAA
TATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAG
TCTCAAGCTTGGCGCGGGGTACCGAGCTCGAATTCCGAGTGTACTTCAAGTCAGTTGGAAATCA
ATAAAATGATTATTTTATGAATATATTTCATTGTGCAAGTAGATAGAAATTACATATGTTACAT
AACACACGAAATAAACAAAAAAACACAATCCAAAACAAACACCCCAAACAAAATAACACTATAT
ATATCCTCGTATGAGGAGAGGCACGTTCAGTGACTCGACGATTCCCGAGCAAAAAAAGTCTCCC
CGTCACACATATAGTGGGTGACGCAATTATCTTCAAAGTAATCCTTCTGTTGACTTGTCATTGA
TAACATCCAGTCTTCGTCAGGATTGCAAAGAATTATAGAAGGGATCCCACCTTTTATTTTCTTC
TTTTTTCCATATTTAGGGTTGACAGTGAAATCAGACTGGCAACCTATTAATTGCTTCCACAATG
GGACGAACTTGAAGGGGATGTCGTCGATGATATTATAGGTGGCGTGTTCATCGTAGTTGGTGAA
GTCGATGGTCCCGTTCCAGTAGTTGTGTCGCCCGAGACTTCTAGCCCAGGTGGTCTTTCCGGTA
CGAGTTGGTCCGCAGATGTAGAGGCTGGGGTGTCTGACCCCAGTCCTTCCCTCATCCTGGTTAG
ATCGGCCATCCACTCAAGGTCAGATTGTGCTTGATCGTAGGAGACAGGATGTATGAAAGTGTAG
GCATCGATGCTTACATGATATAGGTGCGTCTCTCTCCAGTTGTGCAGATCTTCGTGGCAGCGGA
GATCTGATTCTGTGAAGGGCGACACGTACTGCTCAGGTTGTGGAGGAAATAATTTGTTGGCTGA
ATATTCCAGCCATTGAAGCTTTGTTGCCCATTCATGAGGGAATTCTTCTTTGATCATGTCAAGA
TACTCCTCCTTAGACGTTGCAGTCTGGATAATAGTTCGCCATCGTGCGTCAGATTTGCGAGGAG
AGACCTTATGATCTCGGAAATCTCCTCTGGTTTTAATATCTCCGTCCTTTGATATGTAATCAAG
GACTTGTTTAGAGTTTCTAGCTGGCTGGATATTAGGGTGATTTCCTTCAAAATCGAAAAAAGAA
GGATCCCTAATACAAGGTTTTTTATCAAGCTGGATAAGAGCATGATAGTGGGTAGTGCCATCTT
GATGAAGCTCAGAAGCAACACCAAGGAAGAAAATAAGAAAAGGTGTGAGTTTCTCCCAGAGAAA

CTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGGTAGGTAAGGAAAACATATTTAGATTGG
AGTCTGAAGTTCTTGCTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACTCTA
TCTTATAACCGGCGTGGAGGCATGGAGGCAAGGGCATTTTGGTAATTTAAGTAGTTAGTGGAAA
ATGACGTCATTTACTTAAAGACGAAGTCTTGCGACAAGGGGGGCCCACGCCGAATTTTAATATT
ACCGGCGTGGCCCCACCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAA
GTTCCCGCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGATGGTATTTG
ATGGAGCGTATATTGTATCAGGTATTTCCGTCGGATACGAATTATTCGTAC

Fig.5: I-2X35S/CPMV-HT/PDISP/Ha B Brisbane với vòng phân giải protein loại bớt/NOS vào trong hệ khuếch đại BeYDV+Replicaza (Cấu trúc số 1059)

Fig.5A

1039+1059.r (SEQ ID NO: 38)

CTTCCCATCCTCCACCAGGAGGTCTATATTTGGTTCCATTGGCCAGCTTCAA

Fig.5B

1039+1059.c (SEQ ID NO: 39)

CAAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCAC

Fig.5C

Cát xét biểu hiện số 1059 từ BeYDV LIR bên trái sang BeYDV LIR bên phải. PDISP/HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein loại bớt được gạch dưới. (SEQ ID NO:40)

CTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACTCTATCTTATAACCGGCGT
GGAGGCATGGAGGCAAGGGCATTTTGGTAATTTAAGTAGTTAGTGGAAAATGACGTCATTTACT
TAAAGACGAAGTCTTGCGACAAGGGGGGCCCACGCCGAATTTTAATATTACCGGCGTGGCCCCA
CCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAAGTTCCCGCCCACTAG
GGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGATGGTATTTGATAAAGCGTATATTG
TATCAGGTATTTCCGTCGGATACGAATTATTCGTACAAGCTTCTTAAGCCGGTCAACATGGTGG
AGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAAT
TGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGT
CACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAG
GAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAG
GAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAAC

ATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAA
GGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGC
TATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGC
GATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCAC
CCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATG
TGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCT
ATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAA
CGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCT
CTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAAC
GTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAAT
AACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCT
GTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTT
CTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAG
AAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTA
AGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTT
ATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCGAATCTGCACTGGAATA
ACATCGTCAAACTCACCACATGTCGTCAAAACTGCTACTCAAGGGGAGGTCAATGTGACTGGTG
TAATACCACTGACAACAACACCCACCAAATCTCATTTTGCAAATCTCAAAGGAACAGAAACCAG
GGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATCTGGACGTAGCCTTGGGCAGACCAAAA
TGCACGGGGAAAATACCCTCGGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTACATCTG
GGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCCTAACCTTCTCCGAGGATA
CGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAAAATGCACCAGGAGGACCCTAC
AAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTTCGCAACAATGG
CTTGGGCCGTCCCAAAAAACGACAAAAACAAAACAGCAACAAATCCATTAACAATAGAAGTACC
ATACATTTGTACAGAAGGAGAAGACCAAATTACCGTTTGGGGGTTCCACTCTGACAACGAGACC
CAAATGGCAAAGCTCTATGGGGACTCAAAGCCCCAGAAGTTCACCTCATCTGCCAACGGAGTGA
CCACACATTACGTTTCACAGATTGGTGGCTTCCCAAATCAAACAGAAGACGGAGGACTACCACA
AAGTGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGGGAAAACAGGAACAATTACCTAT
CAAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTGCGCAAGTGGCAGGAGCAAGGTAATAAAAG
GATCCTTGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTGGATTAAACAAAAG
CAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAATATGGGTGAAAACA
CCCTTGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTG
CAGGTTGGCACGGATACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAG
CACTCAAGAGGCCATAAACAAGATAACAAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAG
AATCTTCAAAGACTAAGCGGTGCCATGGATGAACTCCACAACGAAATACTAGAACTAGATGAGA
AAGTGGATGATCTCAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAGTCCTGCTTTCCAA
TGAAGGAATAATAAACAGTGAAGATGAACATCTCTTGGCGCTTGAAAGAAAGCTGAAGAAAATG
CTGGGCCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTGAAACCAAACACAAGTGCAACCAGA
CCTGTCTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTCTCCCCACCTTTGA
TTCACTGAATATTACTGCTGCATCTTTAAATGACGATGGATTGGATAATCATACTATACTGCTT
TACTACTCAACTGCTGCCTCCAGTTTGGCTGTAACACTGATGATAGCTATCTTTGTTGTTTATA
TGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGTCTATAAAGGCCTATTTTCTTTAGTTTGAA
TTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTT
ATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACAC
AAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTAT
CGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTC
TTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATG

CATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCG
ATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTAC
TAGATCTCTAGAGTCTCAAGCTTGGCGCGGGGTACCGAGCTCGAATTCCGAGTGTACTTCAAGT
CAGTTGGAAATCAATAAAATGATTATTTTATGAATATATTTCATTGTGCAAGTAGATAGAAATT
ACATATGTTACATAACACACGAAATAAACAAAAAAACACAATCCAAAACAAACACCCCAAACAA
AATAACACTATATATATCCTCGTATGAGGAGAGGCACGTTCAGTGACTCGACGATTCCCGAGCA
AAAAAAGTCTCCCCGTCACACATATAGTGGGTGACGCAATTATCTTCAAAGTAATCCTTCTGTT
GACTTGTCATTGATAACATCCAGTCTTCGTCAGGATTGCAAAGAATTATAGAAGGGATCCCACC
TTTTATTTTCTTCTTTTTTCCATATTTAGGGTTGACAGTGAAATCAGACTGGCAACCTATTAAT
TGCTTCCACAATGGGACGAACTTGAAGGGGATGTCGTCGATGATATTATAGGTGGCGTGTTCAT
CGTAGTTGGTGAAGTCGATGGTCCCGTTCCAGTAGTTGTGTCGCCCGAGACTTCTAGCCCAGGT
GGTCTTTCCGGTACGAGTTGGTCCGCAGATGTAGAGGCTGGGGTGTCTGACCCCAGTCCTTCCC
TCATCCTGGTTAGATCGGCCATCCACTCAAGGTCAGATTGTGCTTGATCGTAGGAGACAGGATG
TATGAAAGTGTAGGCATCGATGCTTACATGATATAGGTGCGTCTCTCTCCAGTTGTGCAGATCT
TCGTGGCAGCGGAGATCTGATTCTGTGAAGGGCGACACGTACTGCTCAGGTTGTGGAGGAAATA
ATTTGTTGGCTGAATATTCCAGCCATTGAAGCTTTGTTGCCCATTCATGAGGGAATTCTTCTTT
GATCATGTCAAGATACTCCTCCTTAGACGTTGCAGTCTGGATAATAGTTCGCCATCGTGCGTCA
GATTTGCGAGGAGAGACCTTATGATCTCGGAAATCTCCTCTGGTTTTAATATCTCCGTCCTTTG
ATATGTAATCAAGGACTTGTTTAGAGTTTCTAGCTGGCTGGATATTAGGGTGATTTCCTTCAAA
ATCGAAAAAAGAAGGATCCCTAATACAAGGTTTTTTATCAAGCTGGATAAGAGCATGATAGTGG
GTAGTGCCATCTTGATGAAGCTCAGAAGCAACACCAAGGAAGAAAATAAGAAAAGGTGTGAGTT
TCTCCCAGAGAAACTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGGTAGGTAAGGAAAAC
ATATTTAGATTGGAGTCTGAAGTTCTTGCTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTT
GCCTCAAACTCTATCTTATAACCGGCGTGGAGGCATGGAGGCAAGGGCATTTTGGTAATTTAAG
TAGTTAGTGGAAAATGACGTCATTTACTTAAAGACGAAGTCTTGCGACAAGGGGGGCCCACGCC
GAATTTTAATATTACCGGCGTGGCCCCACCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGAT
TTAAAGTAGAAAAGTTCCCGCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATACGAT
GTGATGGTATTTGATGGAGCGTATATTGTATCAGGTATTTCCGTCGGATACGAATTATTCGTAC

Fig. 5D

Trình tự axit amin của PDISP/Ha từ cúm B/Brisbane/60/2008 với vòng phân giải protein loại bớt (SEQ ID NO:41)

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSH
FANLKGTETRGKLCPKCLNCTDLDVALGRPKCTGKIPSARVSILHEVRPVTSGCFPIMHDRTKI
RQLPNLLRGYEHIRLSTHNVINAENAPGGPYKIGTSGSCPNITNGNGFFATMAWAVPKNDKNKT
ATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFP
NQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCL
HEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAH
GVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISS
QIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFD
AGEFSLPTFDSLNITAASLNDDGLDNHTILLYYSTAASSLAVTLMIAIFVVYMVSRDNVSCSIC
L

Fig.5E: Trình tự nucleotit của PDISP/HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein loại bớt (SEQ ID NO:43)

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGA
TCTTCGCCGATCGAATCTGCACTGGAATAACATCGTCAAACTCACCACATGTCGTCAAAACTGC
TACTCAAGGGGAGGTCAATGTGACTGGTGTAATACCACTGACAACAACACCCACCAAATCTCAT
TTTGCAAATCTCAAAGGAACAGAAACCAGGGGGAAACTATGCCCAAAATGCCTCAACTGCACAG
ATCTGGACGTAGCCTTGGGCAGACCAAAATGCACGGGGAAAATACCCTCGGCAAGAGTTTCAAT
ACTCCATGAAGTCAGACCTGTTACATCTGGGTGCTTTCCTATAATGCACGACAGAACAAAAATT
AGACAGCTGCCTAACCTTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTATCA
ATGCAGAAAATGCACCAGGAGGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTAC
CAATGGAAACGGATTTTTCGCAACAATGGCTTGGGCCGTCCCAAAAAACGACAAAAACAAAACA
GCAACAAATCCATTAACAATAGAAGTACCATACATTTGTACAGAAGGAGAAGACCAAATTACCG
TTTGGGGGTTCCACTCTGACAACGAGACCCAAATGGCAAAGCTCTATGGGGACTCAAAGCCCCA
GAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCA
AATCAAACAGAAGACGGAGGACTACCACAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAA
AATCTGGGAAAACAGGAACAATTACCTATCAAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTG
CGCAAGTGGCAGGAGCAAGGTAATAAAAGGATCCTTGCCTTTAATTGGAGAAGCAGATTGCCTC
CACGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCA
TAGGAAATTGCCCAATATGGGTGAAAACACCCTTGAAGCTGGCCAATGGAACCAAATATAGACC
TCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCCCATGGGGCACAT
GGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAACAAGATAACAAAAAATC
TCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCCATGGATGAACT
CCACAACGAAATACTAGAACTAGATGAGAAAGTGGATGATCTCAGAGCTGATACAATAAGCTCA
CAAATAGAACTCGCAGTCCTGCTTTCCAATGAAGGAATAATAAACAGTGAAGATGAACATCTCT
TGGCGCTTGAAAGAAAGCTGAAGAAAATGCTGGGCCCCTCTGCTGTAGAGATAGGGAATGGATG
CTTTGAAACCAAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGGTACCTTTGAT
GCAGGAGAATTTTCTCTCCCCACCTTTGATTCACTGAATATTACTGCTGCATCTTTAAATGACG
ATGGATTGGATAATCATACTATACTGCTTTACTACTCAACTGCTGCCTCCAGTTTGGCTGTAAC
ACTGATGATAGCTATCTTTGTTGTTTATATGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGT
CTATAA

Fig. 6: B-2X35S/CPMV-HT/HA B Wisconsin/NOS vào trong hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1462)

Fig.6A, SEQ ID NO:49
IF-HAB110.S1+3c

AAATTTGTCGGGCCCATGAAGGCAATAATTGTACTACTCATGGTAG

Fig.6B: SEQ ID NO:50
IF-HAB110.s1-4r

ACTAAAGAAAATAGGCCTTTATAGACAGATGGAGCATGAAACGTTGTCTCTG

Fig.6C: SEQ ID NO:51
HA B Wisconsin tổng hợp (số truy cập Genbank JN993010)

ATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGGA
TAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGG
CGTGATACCACTGACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACC
AGAGGGAAACTATGCCCGGACTGTCTCAACTGTACAGATCTGGATGTGGCCTTGGGCAGGCCAA
TGTGTGTGGGGACCACACCTTCTGCTAAAGCTTCAATACTCCACGAGGTCAGACCTGTTACATC
CGGGTGCTTTCCTATAATGCACGACAGAACAAAAATCAGGCAACTACCCAATCTTCTCAGAGGA
TATGAAAATATCAGGTTATCAACCCAAAACGTTATCGATGCAGAAAAAGCACCAGGAGGACCCT
ACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAAATCGGATTTTTTGCAACAAT
GGCTTGGGCTGTCCCAAAGGACAACTACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCA
TACATTTGTACAGAAGGGGAAGACCAAATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCC
AAATGAAGAGCCTCTATGGAGACTCAAATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAAC
CACACATTATGTTTCTCAGATTGGCGACTTCCCAGATCAAACAGAAGACGGAGGACTACCACAA
AGCGGCAGAATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTATC
AAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGG
GTCATTGCCTTTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGC
AAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTAAAAACAC
CTTTGAAGCTTGCCAATGGAACCAAATATAGACCTCCTGCAAAACTATTGAAGGAAAGGGGTTT
CTTCGGAGCTATTGCTGGTTTCCTAGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGA
TACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCAGACCTTAAGAGTACACAAGAAGCTA
TAAATAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAACCTTCAAAGACT
AAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTC
AGAGCTGACACTATAAGCTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAA
ACAGTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGC
TGTAGACATAGGAAACGGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTAGACAGG
ATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCATTGAACATTA
CTGCTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGC
TGCTTCTAGTTTGGCTGTAACATTAATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGAC
AACGTTTCATGCTCCATCTGTCTATAA

Fig.6D
Sự biểu diễn theo sơ đồ của cấu trúc 193. Vị trí enzym giới hạn SacII và StuI được sử dụng để tuyến tính hoá plasmit được chú giải đối với sự biểu hiện.

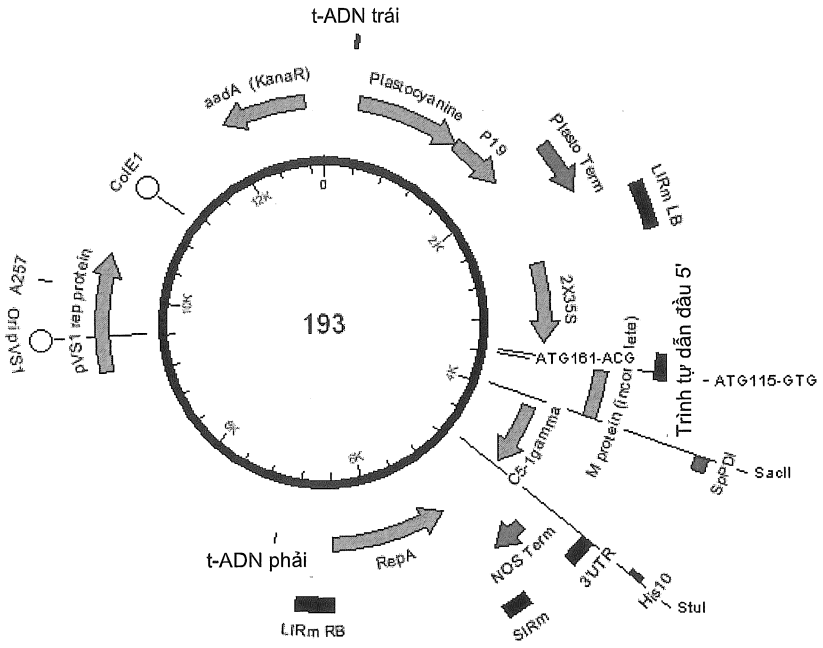


Fig.6E: SEQ ID NO:52
Cấu trúc 193 từ bên trái sang bên phải đường biên t-ADN (gạch dưới). 2X35S/CPMV-HT/NOS vào trong hệ khuếch đại BeYDV(m)+Replicaza với Plastoxyanin-P19-Plastoxyanin làm câm trình tự biểu hiện chất ức chế

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACTTAATAACACATTGCGGACG
TTTTTAATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTC
AAAAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTA
CTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAG
TGATATTTTGACAACAATTTTGTTGCAACATTTGAGAAAATTTTGTTGTTCTCTCTTTTCATTG
GTCAAAAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGA
GAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGC
TACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTA
GAGAATTTTTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTTAAAATTAAAAGTTGAGTCA
TTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAA
TTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCCATAGAGTC
AGTTAACTCATTTTTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAA
AAAAAAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTA
GGAGGATAACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCAC
GCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAA
ACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATT

CCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGA
ACGAGCTATACAAGGAAACGACGCTAGGGAACAAGCTAACAGTGAACGTTGGGATGGAGGATCA
GGAGGTACCACTTCTCCCTTCAAACTTCCTGACGAAAGTCCGAGTTGGACTGAGTGGCGGCTAC
ATAACGATGAGACGAATTCGAATCAAGATAATCCCCTTGGTTTCAAGGAAAGCTGGGGTTTCGG
GAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCACTGCACAGAGTCCTT
GGATCTTGGACGGGAGATTCGGTTAACTATGCAGCATCTCGATTTTTCGGTTTCGACCAGATCG
GATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACTCT
TCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAA
GTGGAAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGT
AAGTTAAAATGCTTCTTCGTCTCCTATTTATAATATGGTTTGTTATTGTTAATTTTGTTCTTGT
AGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTGGTGTAATG
TAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCCATAACTAACTAGACATGAAG
ACCTGCCGCGTACAATTGTCTTATATTTGAACAACTAAAATTGAACATCTTTTGCCACAACTTT
ATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGT
TATCGAAATTCATTAACAATCAACTTAACGTTATTAACTACTAATTTTATATCATCCCCTTTGA
TAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTGTCGTTTCCCGCCTTC
AGTTTGCAAGCTGCTCTAGCCGTGTAGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGGAAT
TACTAGCGCGTGTCGAGACGCGTTGTTGTTGTGACTCCGAGGGGTTGCCTCAAACTCTATCTTA
TAACCGGCGTGGAGGCATGGAGGCAGGGGTATTTTGGTCATTTTAATAGATAGTGGAAAATGAC
GTGGAATTTACTTAAAGACGAAGTCTTTGCGACAAGGGGGGGCCCACGCCGAATTTAATATTAC
CGGCGTGGCCCCCCCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAATT
TCCCGCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGATGGTATTTGGT
CGACAAGCTTGCATGCCGGTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCA
AAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAA
CCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGT
GGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACA
GTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCAC
GTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAA
AATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATAT
CCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAA
GGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCT
GCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTC
CAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACA
ATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTAT
TAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAA
CTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACG
TTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTC
TTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGT
CTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTG
CTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTT
TCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCC
CGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGC
CCGCGGATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTT
CTCAGATCTTCGCCTGCAGGCTCCTCAGCCAAAACGACACCCCCATCTGTCTATCCACTGGCCC
CTGGATCTGCTGCCCAAACTAACTCCATGGTGACCCTGGGATGCCTGGTCAAGGGCTATTTCCC
TGAGCCAGTGACAGTGACCTGGAACTCTGGATCCCTGTCCAGCGGTGTGCACACCTTCCCAGCT
GTCCTGCAGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGTCCCCTCCAGCACCTGGCCCA

GCGAGACCGTCACCTGCAACGTTGCCCACCCGGCCAGCAGCACCAAGGTGGACAAGAAAATTGT
GCCCAGGGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCATC
TTCCCCCCAAAGCCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTGG
TAGACATCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTAGATGATGTGGAGGTGCA
CACAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGCTCAGTCAGTGAACTT
CCCATCATGCACCAGGACTGGCTCAATGGCAAGGAGCGATCGCTCACCATCACCATCACCATCA
CCATCACCATTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTAT
GTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAG
CTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAAAAAAAA
AAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGC
AATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTT
GAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTT
ATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACT
AGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGC
GCCATAAAATGATTATTTTATGAATATATTTCATTGTGCAAGTAGATAGAAATTACATATGTTA
CATAACACACGAAATAAACAAAAAAAGACAATCCAAAAACAAACACCCCAAAAAAAATAATCAC
TTTAGATAAACTCGTATGAGGAGAGGCACGTTCAGTGACTCGACGATTCCCGAGCAAAAAAAGT
CTCCCCGTCACACATATAGTGGGTGACGCAATTATCTTTAAAGTAATCCTTCTGTTGACTTGTC
ATTGATAACATCCAGTCTTCGTCAGGATTGCAAAGAATTATAGAAGGGATCCCACCTTTTATTT
TCTTCTTTTTTCCATATTTAGGGTTGACAGTGAAATCAGACTGGCAACCTATTAATTGCTTCCA
CAATGGGACGAACTTGAAGGGGATGTCGTCGATGATATTATAGGTGGCGTGTTCATCGTAGTTG
GTGAAATCGATGGTACCGTTCCAATAGTTGTGTCGTCCGAGACTTCTAGCCCAGGTGGTCTTTC
CGGTACGAGTTGGTCCGCAGATGTAGAGGCTGGGGTGTCGGATTCCATTCCTTCCATTGTCCTG
GTTAAATCGGCCATCCATTCAAGGTCAGATTGAGCTTGTTGGTATGAGACAGGATGTATGTAAG
TATAAGCGTCTATGCTTACATGGTATAGATGGGTTTCCCTCCAGGAGTGTAGATCTTCGTGGCA
GCGAAGATCTGATTCTGTGAAGGGCGACACATACGGTTCAGGTTGTGGAGGGAATAATTTGTTG
GCTGAATATTCCAGCCATTGAAGTTTTGTTGCCCATTCATGAGGGAATTCTTCCTTGATCATGT
CAAGATATTCCTCCTTAGACGTTGCAGTCTGGATAATAGTTCTCCATCGTGCGTCAGATTTGCG
AGGAGAGACCTTATGATCTCGGAAATCTCCTCTGGTTTTAATATCTCCGTCCTTTGATATGTAA
TCAAGGACTTGTTTAGAGTTTCTAGCTGGCTGGATATTAGGGTGATTTCCTTCAAAATCGAAAA
AAGAAGGATCCCTAATACAAGGTTTTTTATCAAGCTGGAGAAGAGCATGATAGTGGGTAGTGCC
ATCTTGATGAAGCTCAGAAGCAACACCAAGGAAGAAAATAAGAAAAGGTGTGAGTTTCTCCCAG
AGAAACTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGATAGGTAAGGAAAACATATTTAG
ATTGGAGTCTGAAGTTCTTACTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAA
CTCTATCTTATAACCGGCGTGGAGGCATGGAGGCAGGGGTATTTTGGTCATTTTAATAGATAGT
GGAAAATGACGTGGAATTTACTTAAAGACGAAGTCTTTGCGACAAGGGGGGGCCCACGCCGAAT
TTAATATTACCGGCGTGGCCCCCCCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAA
GTAGAAAATTTCCCGCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGAT
GGTATTTGACTAGTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTA
CCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCG
CACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTG
GATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGACAGGATATATTGGCGGG
TAAACCTAAGAGAAAAGAGCGTTTA

Fig.6F: SEQ ID NO:53
Cát xét biểu hiện số 1462 từ đoạn khởi đầu 2X35S đến kết thúc NOS. HA từ cúm B/Wisconsin/1/2010 được gạch dưới.

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAG
ACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCAT
CATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGAC
CCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGA
TTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTC
TCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGAT
TCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAA
ATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAA
GATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGC
AAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCA
AGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTT
TGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTT
AAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCG
GCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGC
GGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTT
GCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGC
AATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGA
TTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGG
AGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGAAGGCAATAATTG
TACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGGATAACATCTTCAAACTC
ACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGCGTGATACCACTGACA
ACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGCC
CGGACTGTCTCAACTGTACAGATCTGGATGTGGCCTTGGGCAGGCCAATGTGTGTGGGGACCAC
ACCTTCTGCTAAAGCTTCAATACTCCACGAGGTCAGACCTGTTACATCCGGGTGCTTTCCTATA
ATGCACGACAGAACAAAAATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAAATATCAGGT
TATCAACCCAAAACGTTATCGATGCAGAAAAAGCACCAGGAGGACCCTACAGACTTGGAACCTC
AGGATCTTGCCCTAACGCTACCAGTAAAATCGGATTTTTTGCAACAATGGCTTGGGCTGTCCCA
AAGGACAACTACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAAG
GGGAAGACCAAATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAGCCTCTA
TGGAGACTCAAATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACATTATGTTTCT
CAGATTGGCGACTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAGCGGCAGAATTGTTG
TTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTATCAAAGAGGTGTTTTGTT
GCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCATTGCCTTTAATT
GGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAG
GAGAACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTAAAAACACCTTTGAAGCTTGCCAA
TGGAACCAAATATAGACCTCCTGCAAAACTATTGAAGGAAAGGGGTTTCTTCGGAGCTATTGCT
GGTTTCCTAGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCTCACGGAG
CACATGGAGTGGCAGTGGCGGCAGACCTTAAGAGTACACAAGAAGCTATAAATAAGATAACAAA
AAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAACCTTCAAAGACTAAGTGGTGCCATGGAT
GAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAA
GCTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGACGAGCA
TCTATTGGCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGACATAGGAAAC

GGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCT
TTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCATTGAACATTACTGCTGCATCTTTAAA
TGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTTCTAGTTTGGCT
GTAACATTAATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCCA
TCTGTCTATAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGT
TTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCT
CCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAAAAAAAAAA
AAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAA
TAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGA
ATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTAT
GATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAG
GATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.6G: SEQ ID NO:54
Trình tự axit amin của HA từ cúm B/Wisconsin/1/2010

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTRT
RGKLCPDCLNCTDLDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLPNLLRG
YENIRLSTQNVIDAEKAPGGPYRLGTSGSCPNATSKIGFFATMAWAVPKDNYKNATNPLTVEVP
YICTEGEDQITVWGFHSDNKTQMKSLYGDSNPQKFTSSANGVTTHYVSQIGDFPDQTEDGGLPQ
SGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKS
KPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHG
YTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDL
RADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVDIGNGCFETKHKCNQTCLDR
IAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNHTILLYYSTAASSLAVTLMLAIFIVYMVSRD
NVSCSICL*

Fig.6H
Sự biểu diễn theo sơ đồ của cấu trúc số 1462

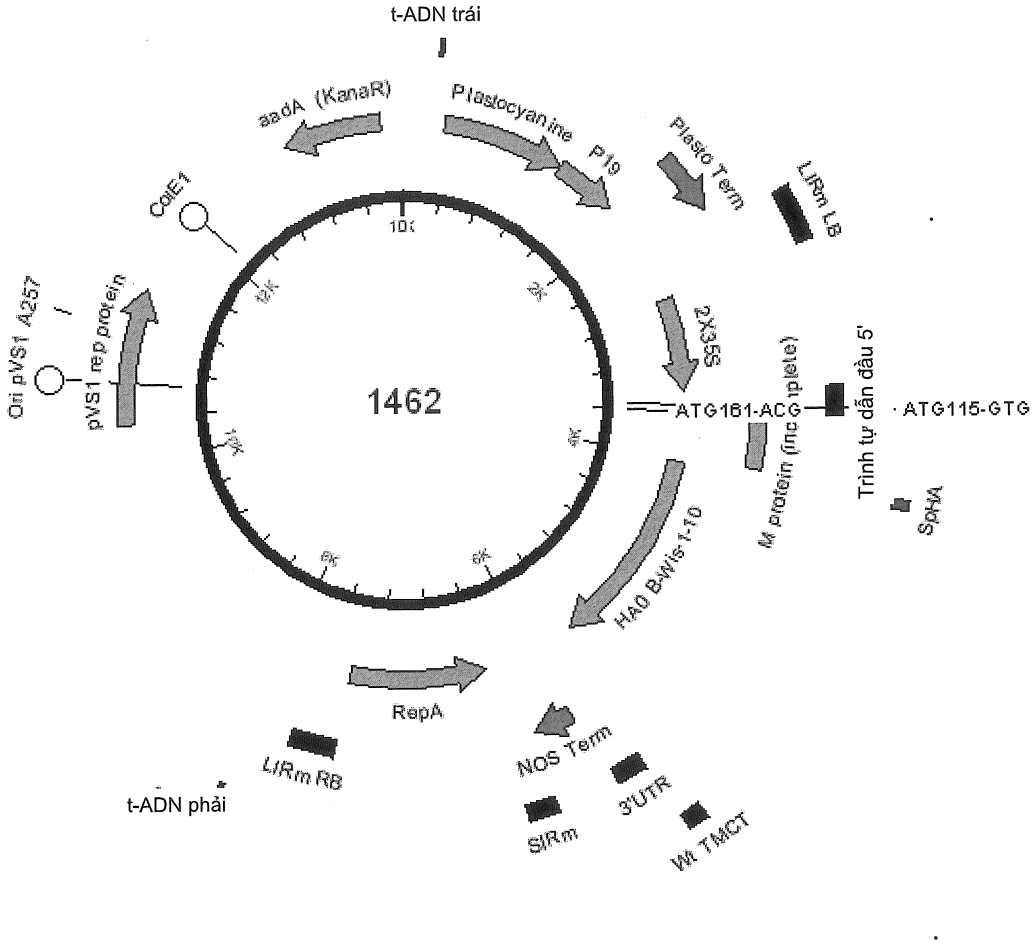


Fig.7: C-2X35S/CPMV-HT/HA B Wisconsin với vòng phân giải protein loại bớt/NOS vào trong hệ khuếch đại BeYDV(m)+Replicaza (Cấu trúc số 1467)

Fig.7A: SEQ ID NO:55
HAB110(PrL-).r

TCCTTCCCATCCTCCACCAGGAGGTCTATATTTGGTTCCATTGGCAAGCTTCAAAG

Fig.7B: SEQ ID NO:56
HAB110(PrL-).c

ATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGA

Fig.7C: SEQ ID NO:57
Cát xét biểu hiện số 1467 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. HA từ cúm B/Wisconsin/1/2010 với vòng phân giải protein loại bớt được gạch dưới.

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAG
ACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCAT
CATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGAC
CCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGA
TTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTC
TCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGAT
TCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAA
ATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAA
GATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGC
AAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCA
AGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTT
TGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTT
AAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCG
GCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGC
GGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTT
GCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGC
AATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGA
TTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGG
AGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGAAGGCAATAATTG
TACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGGATAACATCTTCAAACTC
ACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGCGTGATACCACTGACA
ACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGCC
CGGACTGTCTCAACTGTACAGATCTGGATGTGGCCTTGGGCAGGCCAATGTGTGTGGGGACCAC
ACCTTCTGCTAAAGCTTCAATACTCCACGAGGTCAGACCTGTTACATCCGGGTGCTTTCCTATA
ATGCACGACAGAACAAAAATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAAATATCAGGT
TATCAACCCAAAACGTTATCGATGCAGAAAAAGCACCAGGAGGACCCTACAGACTTGGAACCTC
AGGATCTTGCCCTAACGCTACCAGTAAAATCGGATTTTTTGCAACAATGGCTTGGGCTGTCCCA
AAGGACAACTACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAAG
GGGAAGACCAAATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAGCCTCTA

TGGAGACTCAAATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACATTATGTTTCT
CAGATTGGCGACTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAGCGGCAGAATTGTTG
TTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTATCAAAGAGGTGTTTTGTT
GCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCATTGCCTTTAATT
GGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAG
GAGAACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTAAAAACACCTTTGAAGCTTGCCAA
TGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATAC
ACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCAGACCTTAAGAGTACACAAGAAGCTATAA
ATAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAACCTTCAAAGACTAAG
TGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGA
GCTGACACTATAAGCTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACA
GTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGT
AGACATAGGAAACGGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTAGACAGGATA
GCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCATTGAACATTACTG
CTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGC
TTCTAGTTTGGCTGTAACATTAATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAAC
GTTTCATGCTCCATCTGTCTATAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTG
TGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAAT
TTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTT
ATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTC
AAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATA
TAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGA
GATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATA
GCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.7D: SEQ ID NO:58
Trình tự axit amin của HA từ cúm B/Wisconsin/1/2010 với vòng phân giải protein loại bớt

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTRT
RGKLCPDCLNCTDLDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLPNLLRG
YENIRLSTQNVIDAEKAPGGPYRLGTSGSCPNATSKIGFFATMAWAVPKDNYKNATNPLTVEVP
YICTEGEDQITVWGFHSDNKTQMKSLYGDSNPQKFTSSANGVTTHYVSQIGDFPDQTEDGGLPQ
SGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKS
KPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKS
TQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSN
EGIINSEDEHLLALERKLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFD
SLNITAASLNDDGLDNHTILLYYSTAASSLAVTLMLAIFIVYMVSRDNVSCSICL*

Fig.7E
Sự biểu diễn theo sơ đồ của cấu trúc số 1467

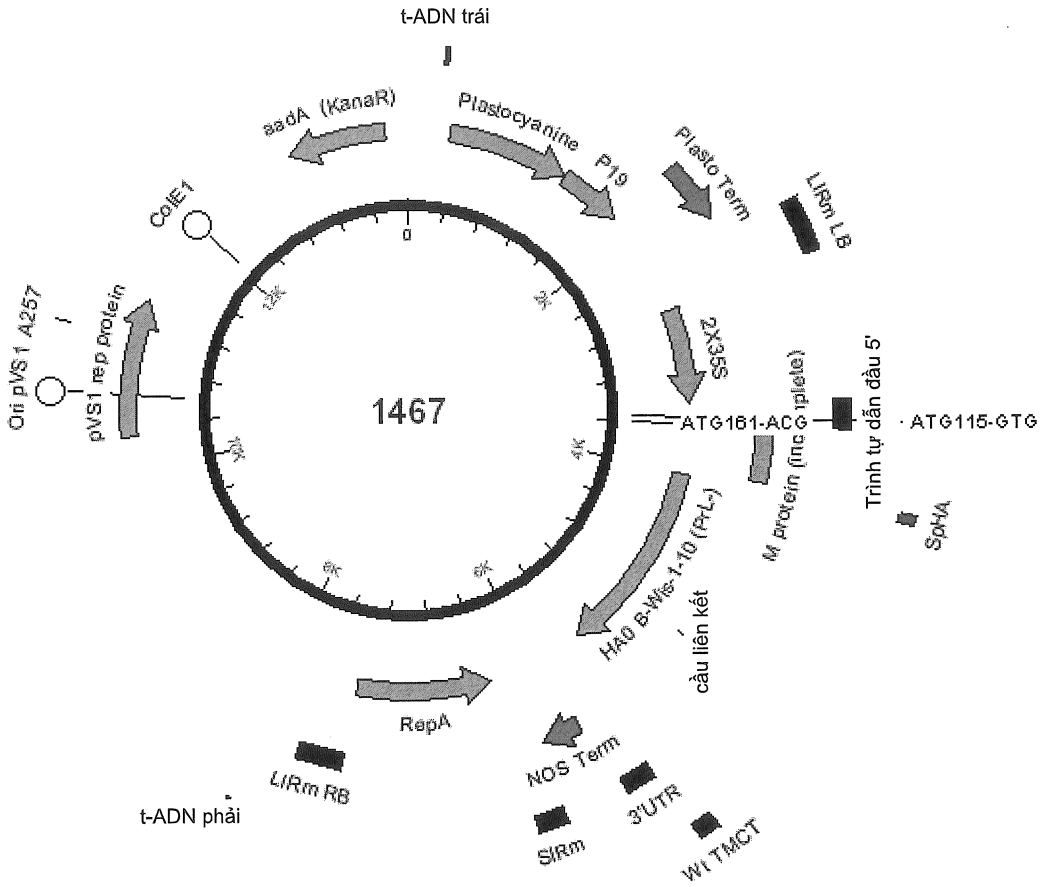


Fig.8: A-2X35S/CPMV-HT/PDISP/HA B Brisbane với vòng phân giải protein loại bớt/NOS (Cấu trúc số 1039)

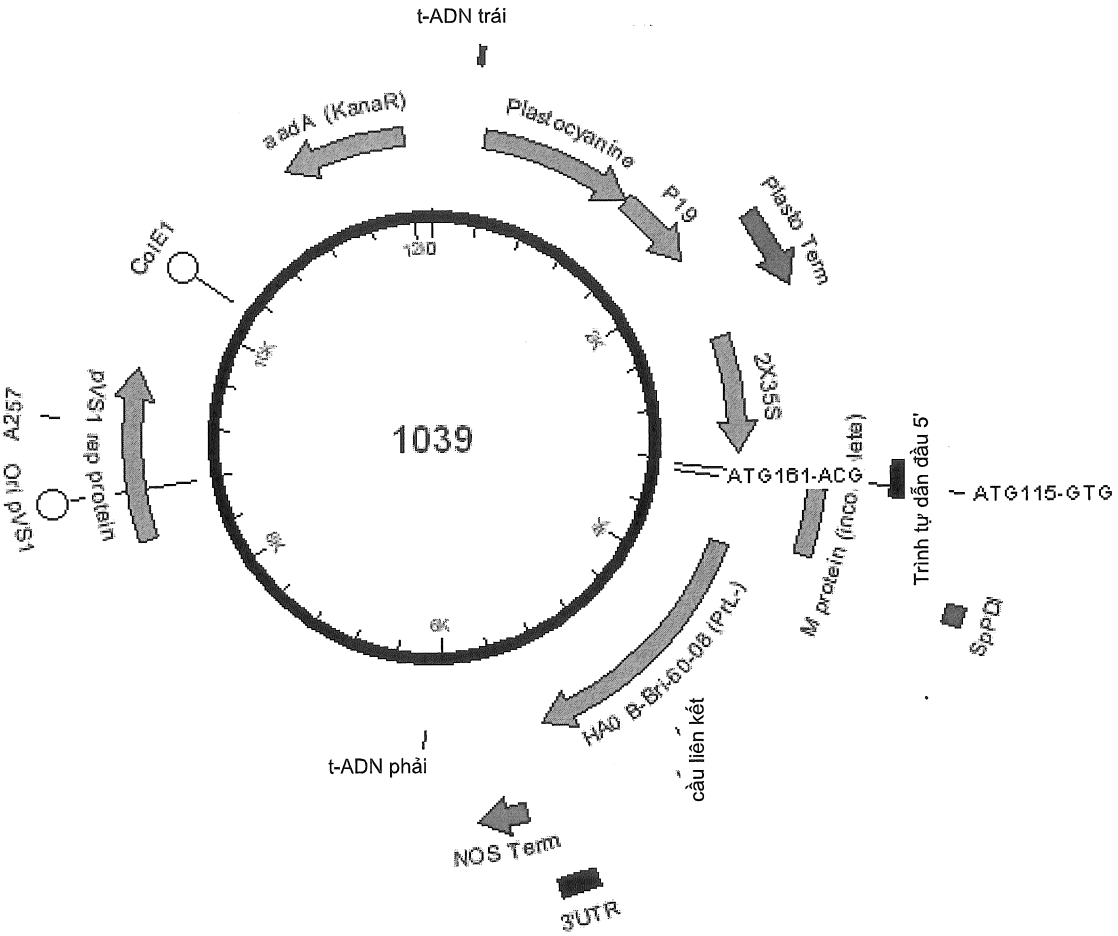
Fig.8A (SEQ ID NO:15)

Cát xét biểu hiện số 1039 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. PDISP/HA từ cúm B/Brisbane/60/2008 với vòng phân giải protein loại bớt được gạch dưới

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAG
ACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCAT
CATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGAC
CCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGA
TTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTC
TCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGAT
TCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAA
ATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAA
GATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGC
AAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCA
AGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTT
TGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTT
AAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCG
GCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGC
GGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTT
GCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGC
AATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGA
TTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGG
AGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGGCGAAAAACGTTG
CGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCGAAT
CTGCACTGGAATAACATCGTCAAACTCACCACATGTCGTCAAAACTGCTACTCAAGGGGAGGTC
AATGTGACTGGTGTAATACCACTGACAACAACACCCACCAAATCTCATTTTGCAAATCTCAAAG
GAACAGAAACCAGGGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATCTGGACGTAGCCTT
GGGCAGACCAAAATGCACGGGGAAAATACCCTCGGCAAGAGTTTCAATACTCCATGAAGTCAGA
CCTGTTACATCTGGGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCCTAACC
TTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAAAATGCACC
AGGAGGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTT
TTCGCAACAATGGCTTGGGCCGTCCCAAAAAACGACAAAAACAAAACAGCAACAAATCCATTAA
CAATAGAAGTACCATACATTTGTACAGAAGGAGAAGACCAAATTACCGTTTGGGGGTTCCACTC
TGACAACGAGACCCAAATGGCAAAGCTCTATGGGGACTCAAAGCCCCAGAAGTTCACCTCATCT
GCCAACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCAAATCAAACAGAAGACG
GAGGACTACCACAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGGGAAAACAGG
AACAATTACCTATCAAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTGCGCAAGTGGCAGGAGC
AAGGTAATAAAAGGATCCTTGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTG
GATTAAACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAAT
ATGGGTGAAAACACCCTTGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGG
GAAGGAATGATTGCAGGTTGGCACGGATACACATCCCATGGGGCACATGGAGTAGCGGTGGCAG
CAGACCTTAAGAGCACTCAAGAGGCCATAAACAAGATAACAAAAAATCTCAACTCTTTGAGTGA

GCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCCATGGATGAACTCCACAACGAAATACTA
GAACTAGATGAGAAAGTGGATGATCTCAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAG
TCCTGCTTTCCAATGAAGGAATAATAAACAGTGAAGATGAACATCTCTTGGCGCTTGAAAGAAA
GCTGAAGAAAATGCTGGGCCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTGAAACCAAACAC
AAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTC
TCCCCACCTTTGATTCACTGAATATTACTGCTGCATCTTTAAATGACGATGGATTGGATAATCA
TACTATACTGCTTTACTACTCAACTGCTGCCTCCAGTTTGGCTGTAACACTGATGATAGCTATC
TTTGTTGTTTATATGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGTCTATAAAGGCCTATTT
TCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGC
TCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCT
TCAGCAAGGACACAAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAAGACCGGGAATTCG
ATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAAT
CCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAA
TTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATA
CATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTG
TCATCTATGTTACTAGAT

Fig.8B
Sự biểu diễn theo sơ đồ của cấu trúc số 1039



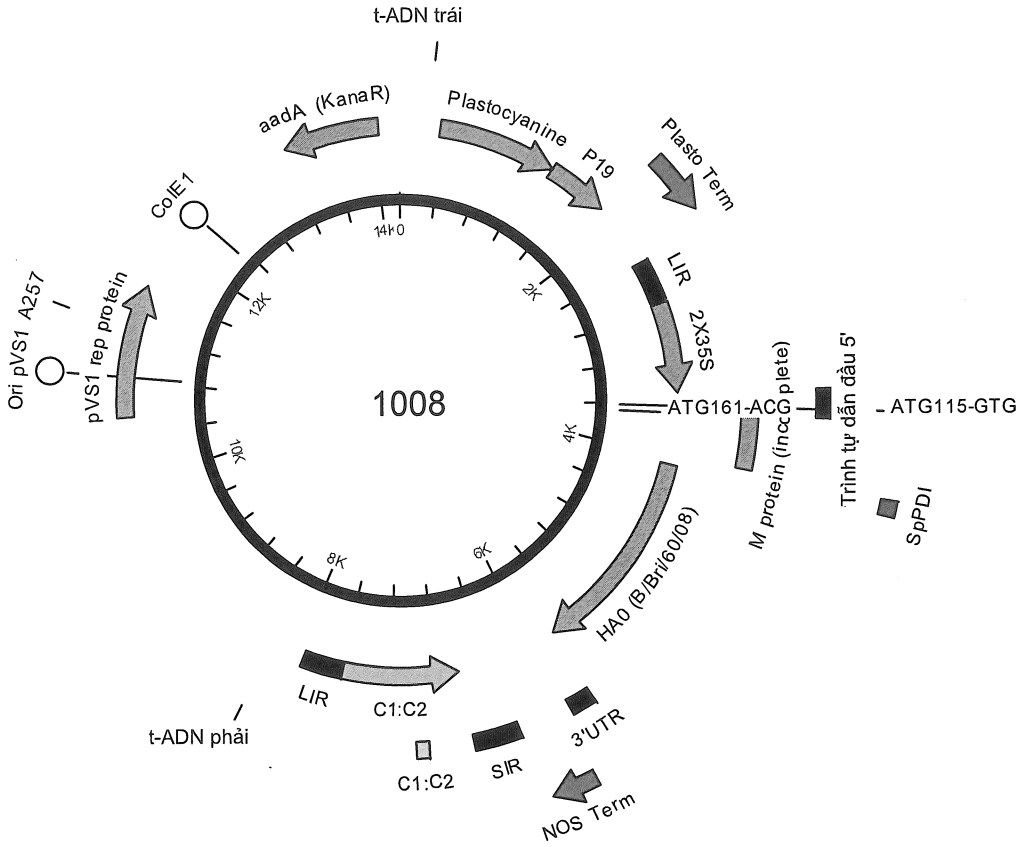

t-ADN trái
aadA (KanaR)
Plastocyanine P19
Plasto Term
ColE1
LIR 2X35S
Ori pVS1 A257
pVS1 rep protein
1008
ATG161-ACG plete
ATG115-GTG
M protein (inc
Trình tự dẫn đầu 5'
SpPDI
HA0 (B/Bri/60/08)
LIR
C1:C2
3'UTR
t-ADN phải
C1:C2
SIR
NOS Term

Fig.9

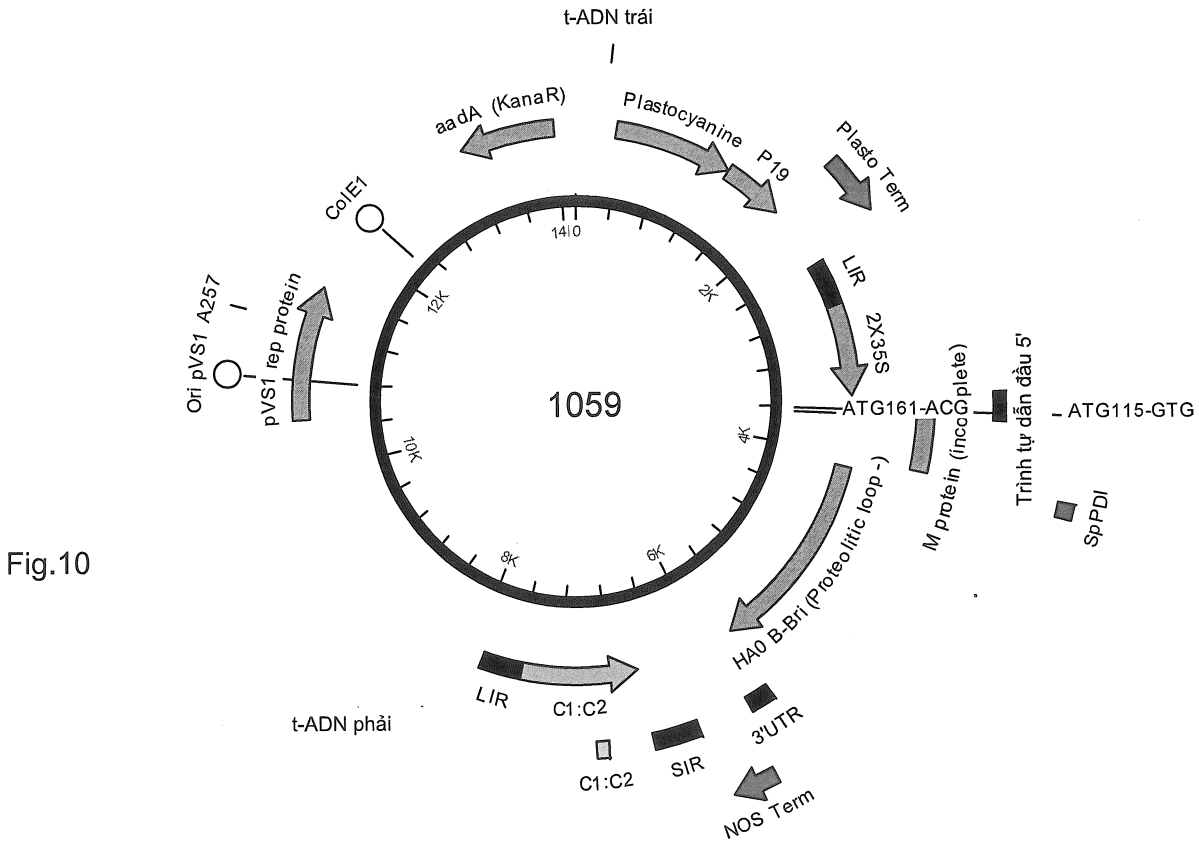

t-ADN trái
aadA (KanaR)
Plastocyanine P19
Plasto Term
ColE1
LIR 2X35S
Ori pVS1 A257
pVS1 rep protein
1059
ATG161-ACG plete
ATG115-GTG
M protein (inc
Trình tự dẫn đầu 5'
SpPDI
HA0 B-Bri (Proteolitic loop -)
LIR
C1:C2
3'UTR
t-ADN phải
C1:C2
SIR
NOS Term

Fig.10

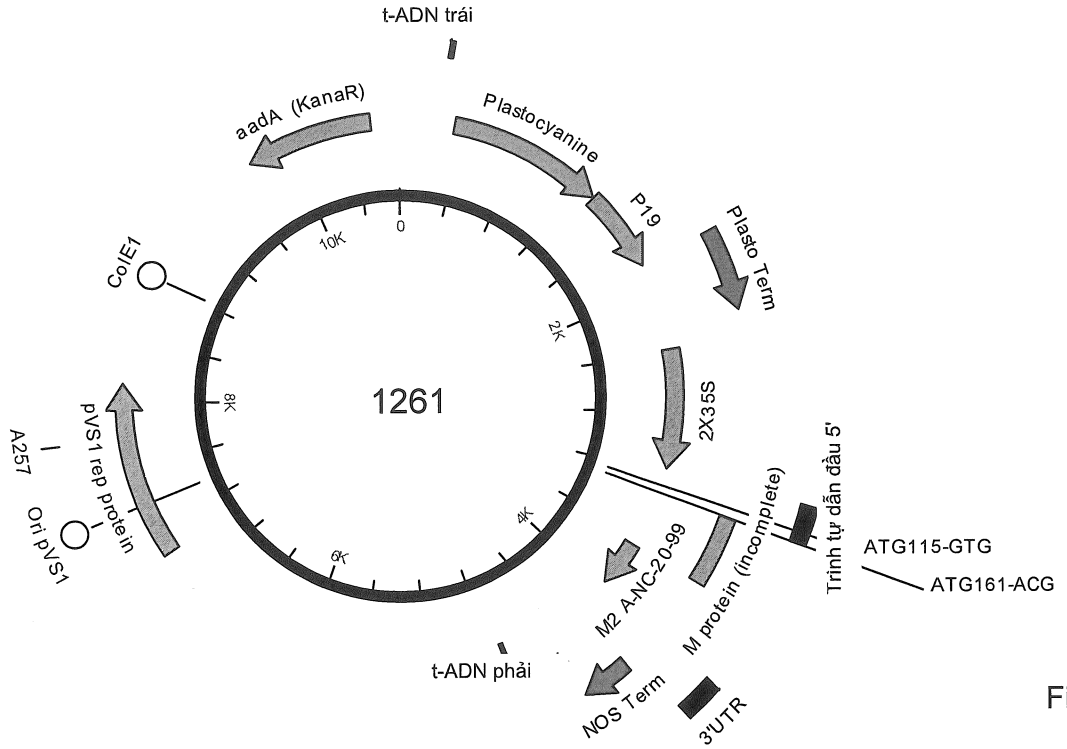

t-ADN trái
aadA (KanaR)
Plastocyanine
P19
Plasto Term
2X35S
ATG115-GTG
ATG161-ACG
Trình tự dẫn đầu 5'
M protein (incomplete)
M2 A-NC-20-99
t-ADN phải
NOS Term
3'UTR
CoIE1
pVS1 rep protein
Ori pVS1
A257
1261
0
2K
4K
6K
8K
10K

Fig.11

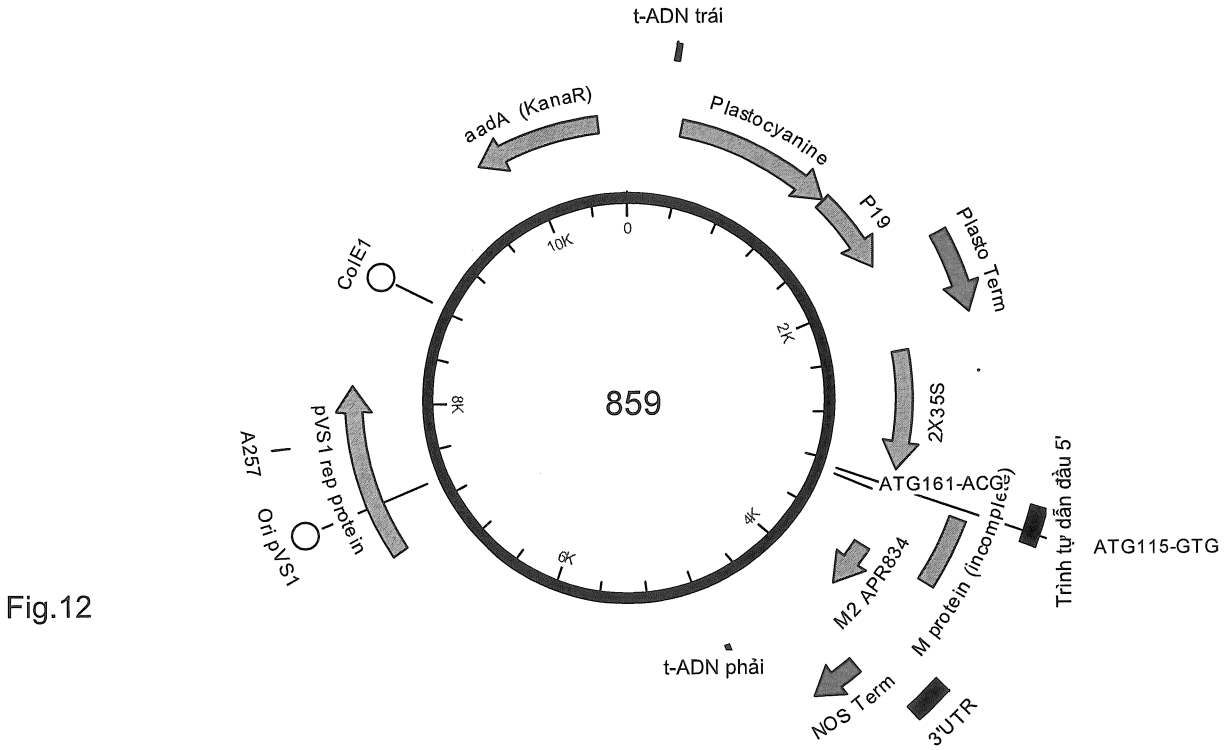

t-ADN trái
aadA (KanaR)
Plastocyanine
P19
Plasto Term
2X35S
ATG161-ACG
ATG115-GTG
Trình tự dẫn đầu 5'
M protein (incomplete)
M2 APR834
t-ADN phải
NOS Term
3'UTR
CoIE1
pVS1 rep protein
Ori pVS1
A257
859
0
2K
4K
6K
8K
10K

Fig.12

Fig.13A

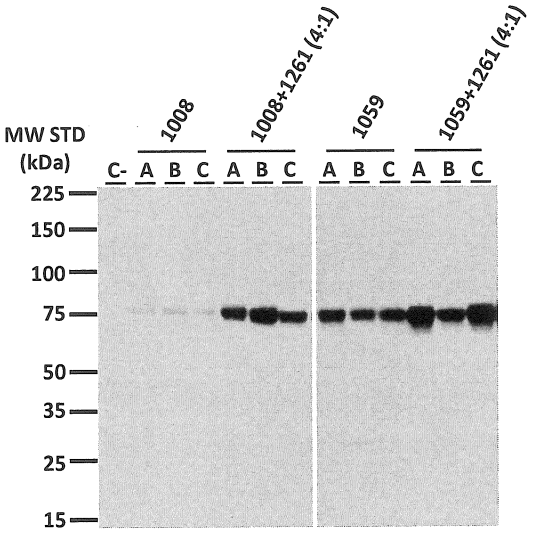


Fig.13B

Cấu trúc số	Khả năng hemagglutin hoá (đơn vị HA/mg protein)a
1008	533
1008 + 1261 (4:1)	3200
1059	4267
1059 + 1261 (4:1)	34133

[a] Nghịch đảo của lượng nhỏ nhất của tổng protein được đòi hỏi đối với phản ứng hemagglutin hoá với thể tích cuối cùng 200µl.

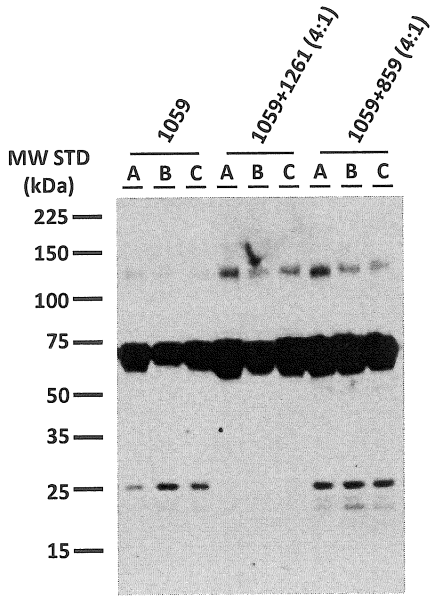

1059
1059+1261 (4:1)
1059+859 (4:1)
MW STD
(kDa)
A B C A B C A B C
225
150
100
75
50
35
25
15

Fig.14

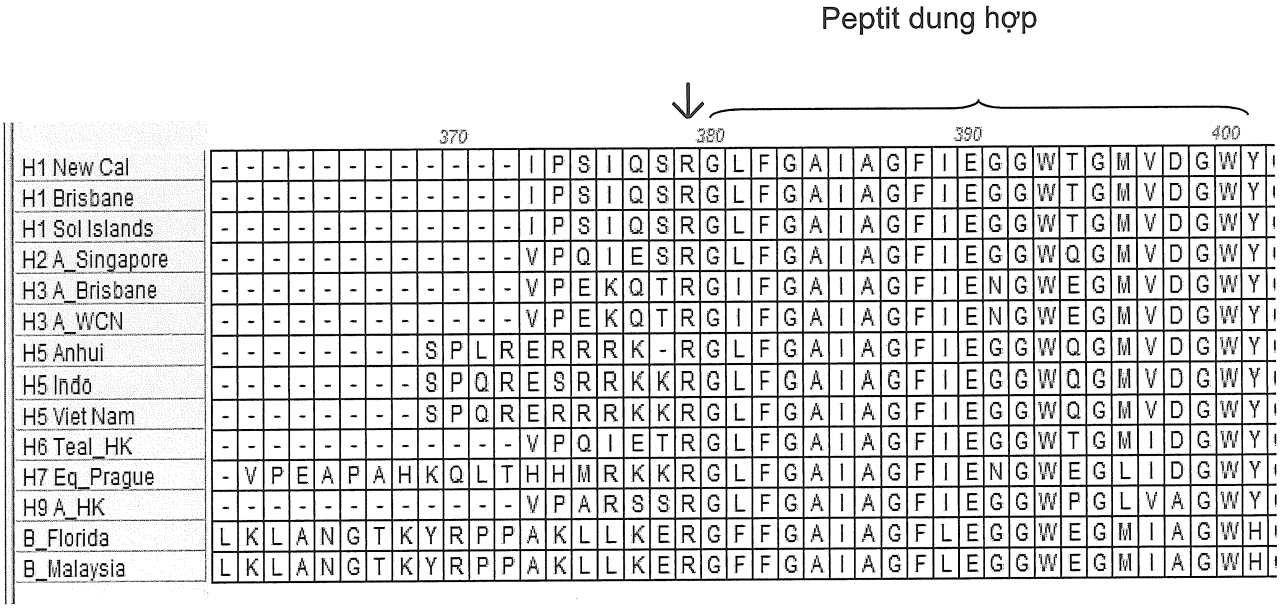

Peptit dung hợp
370
380
390
400
H1 New Cal
H1 Brisbane
H1 Sol Islands
H2 A_Singapore
H3 A_Brisbane
H3 A_WCN
H5 Anhui
H5 Indo
H5 Viet Nam
H6 Teal_HK
H7 Eq_Prague
H9 A_HK
B_Florida
B_Malaysia

Fig.15

Fig.16A

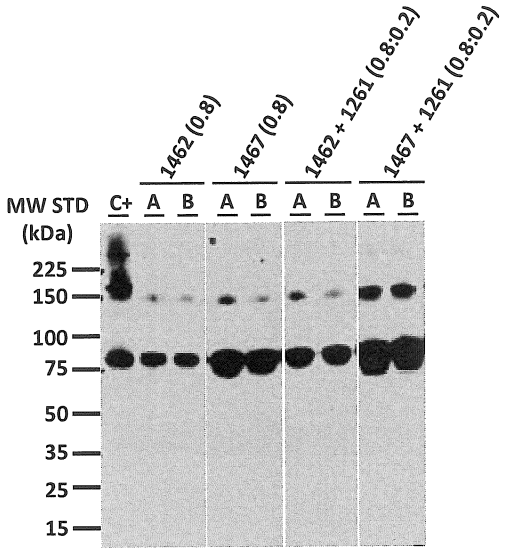


Fig.16B

Cấu trúc số	Khả năng hemagglutin hoá (độ chuẩn HA)a
1462 (0.8)	136 ± 48
1467 (0.8)	2660 ± 384
1462 + 1261 (0.8:0.2)	941 ± 192
1467 + 1261 (0.8:0.2)	4344 ± 1536

[a] Sự nghịch đảo về khả năng pha loãng cao nhất của hồng huyết cầu agglutin hoá.

Fig.17A

Cấu trúc số	Yếu tố điều hoà	BeYDV	Vòng phân giải protein	Đồng biểu hiện M2	Khả năng hemagglutin hoá (đơn vị HA/ml)
1008	CPMV-HT	Có	Có mặt	Không	1216
1008 + 1261	CPMV-HT	Có	Có mặt	Có	5696
1039	CPMV-HT	Không	Loại bỏ	Không	141
1039 + 1261	CPMV-HT	Không	Loại bỏ	Có	2944
1059	CPMV-HT	Có	Loại bỏ	Không	10197
1059 + 1261	CPMV-HT	Có	Loại bỏ	Có	43008

Fig.17B

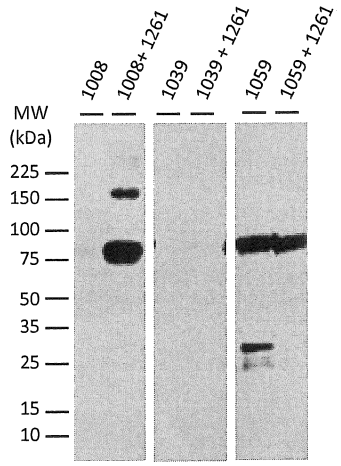


Fig. 18A

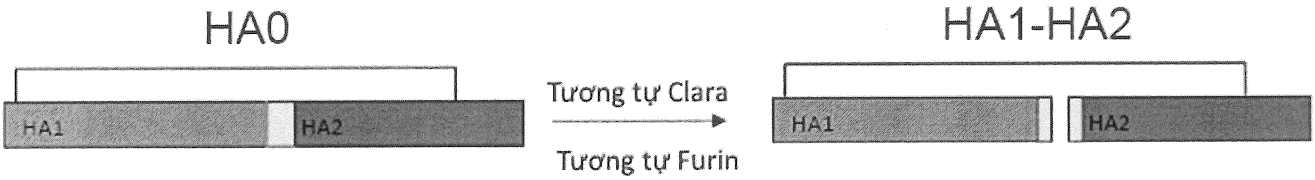


Fig. 18B

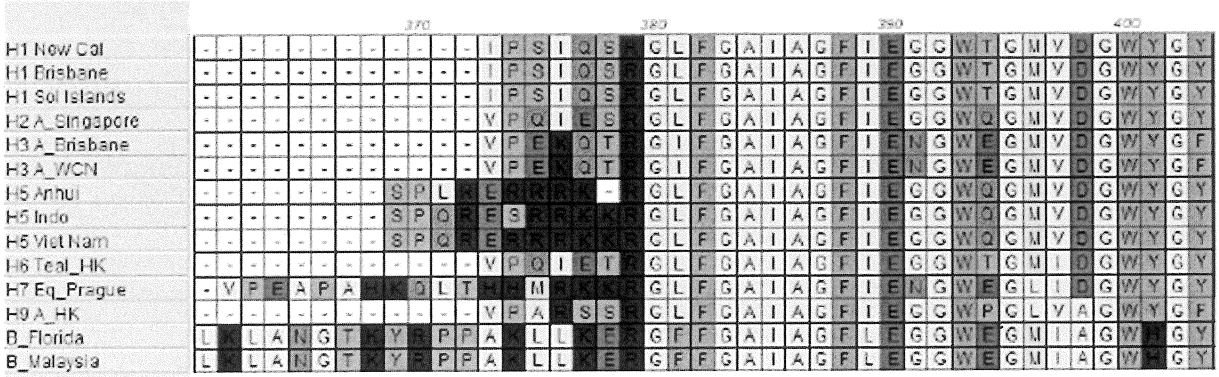


Fig. 18C

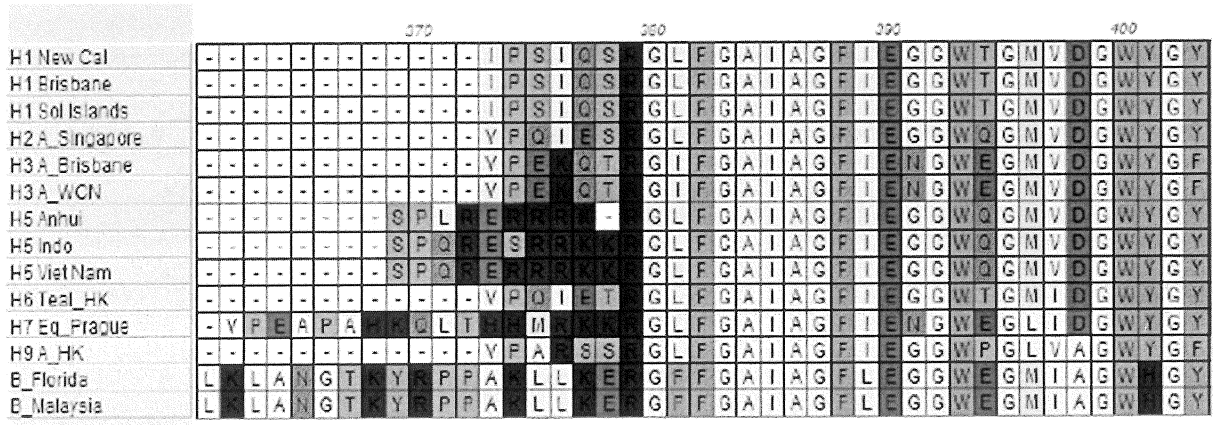
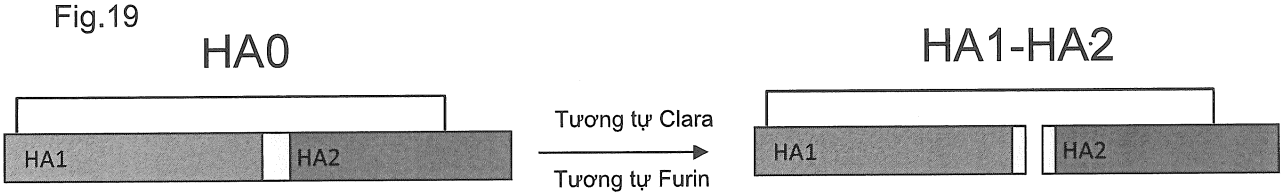


Fig.19



Chủng	Trình tự tự nhiên	Huỷ bỏ vị trí tách tiền chất
H5/Indo	TGLRNSPQ**RESRRKKR**↓GLF	TGLRNSPQTETR↓GLF TGLRNSPQTETQGLF

Fig.20

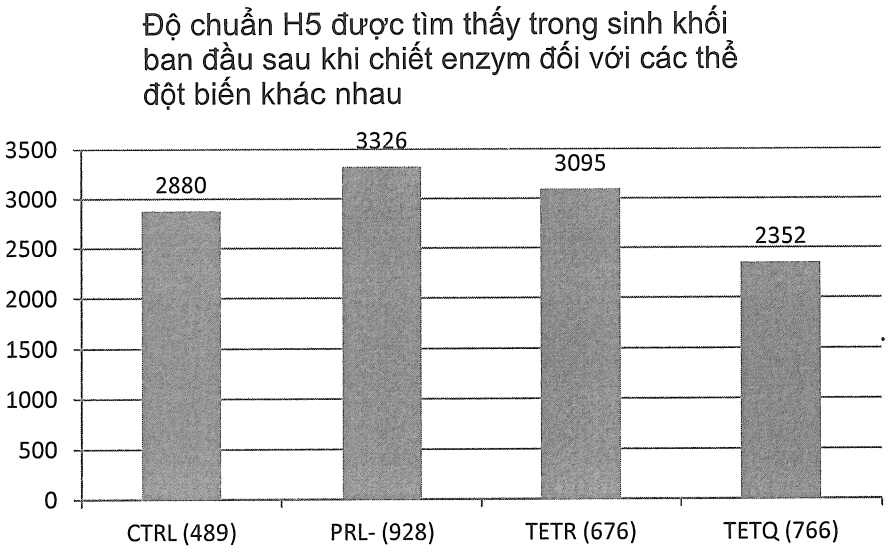

Độ chuẩn H5 được tìm thấy trong sinh khối ban đầu sau khi chiết enzym đối với các thể đột biến khác nhau
3500
3000
2500
2000
1500
1000
500
0
2880
3326
3095
2352
CTRL (489)
PRL- (928)
TETR (676)
TETQ (766)

Fig.21: Phương pháp loại bớt/thay thế đối với typ B HA

Fig.21A

Trình sự sơ cấp của B/Brisbane/60/2008 nguyên thể (SEQ ID NO:16)

DRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKGTETRGKLCPKCLNCTDLD
VALGRPKCTGKIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAE
NAPGGPYKIGTSGSCPNITNGNGFFATMAWAVPKNDKNKTATNPLTIEVPYICTEGEDQITVWG
FHSDNETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSG
KTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGN
CPIWVKTPLKLANGTKYRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAAD
LKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVL
LSNEGIINSEDEHLLALERKLKKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLP
TFDSLNITAASLNDDGLDNHTILLYYSTAASSLAVTLMIAIFVVYMVSRDNVSCSICL

Fig.21B

Δ vòng phân giải protein typ B HA (SEQ ID NO:17)

DRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKGTETRGKLCPKCLNCTDLD
VALGRPKCTGKIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAE
NAPGGPYKIGTSGSCPNITNGNGFFATMAWAVPKNDKNKTATNPLTIEVPYICTEGEDQITVWG
FHSDNETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSG
KTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGN
CPIWVKTPLKLANGTKYRPP-*GG*-
GWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNE
ILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVEIGNGCFET
KHKCNQTCLDRIAAGTFDAGEFSLPTFDSLNITAASLNDDGLDNHTILLYYSTAASSLAVTLMI
AIFVVYMVSRDNVSCSICL

Fig.21C

Sự thay thế vị trí tách của typ B HA bằng cầu liên kết (SEQ ID NO:18)

DRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKGTETRGKLCPKCLNCTDLD
VALGRPKCTGKIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAE
NAPGGPYKIGTSGSCPNITNGNGFFATMAWAVPKNDKNKTATNPLTIEVPYICTEGEDQITVWG
FHSDNETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSG
KTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGN
CPIWVKTPLKLANGTKYR *-GSSSGSSSG-*
GFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQ
RLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGP
SAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSLNITAASLNDDGLDNHTILLYYS
TAASSLAVTLMIAIFVVYMVSRDNVSCSICL

Fig.21D
Phương pháp loại bớt/thay thế đối với H3

Trình tự sơ cấp của H3 A/Perth/16/2009 nguyên thể (SEQ ID NO:19)

QKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGEICDSPHQILDGKNC
TLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFNNESFNWT
GVTQNGTSSACIRRSKNSFFSRLNWLTHLNFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKDQI
FLYAQASGRITVSTKRSQQTVSPNIGSRPRVRNIPSRISIYWTIVKPGDILLINSTGNLIAPRG
YFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRITYGACPRYVKQNTLKLATGM
RNVPEKQTRGIFGAIAGFIENGWEGMVDGWYGFRHQNSEGRGQAADLKSTQAAIDQINGKLNRL
IGKTNEKFHQIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLF
EKTKKQLRENAEDMGNGCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYK
DWILWISFAISCFLLCVALLGFIMWACQKGNIRCNICI

Fig.21E

Δ vòng phân giải protein H3 (trong BeYDV, 1096) (SEQ ID NO:20)

QKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGEICDSPHQILDGKNC
TLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFNNESFNWT
GVTQNGTSSACIRRSKNSFFSRLNWLTHLNFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKDQI
FLYAQASGRITVSTKRSQQTVSPNIGSRPRVRNIPSRISIYWTIVKPGDILLINSTGNLIAPRG
YFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRITYGACPRYVKQNTLKLATGM
-GS-
GAIAGFIENGWEGMVDGWYGFRHQNSEGRGQAADLKSTQAAIDQINGKLNRLIGKTNEKFHQIE
KEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEKTKKQLRENAE
DMGNGCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWISFAISC
FLLCVALLGFIMWACQKGNIRCNICI

Fig.21F

Sự thay thế vị trí tách của H3 bằng cầu liên kết (trong BeYDV, 1097) (SEQ ID NO:21)

QKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGEICDSPHQILDGKNC
TLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFNNESFNWT
GVTQNGTSSACIRRSKNSFFSRLNWLTHLNFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKDQI
FLYAQASGRITVSTKRSQQTVSPNIGSRPRVRNIPSRISIYWTIVKPGDILLINSTGNLIAPRG
YFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRITYGACPRYVKQNTLKLATGM
-GSSGSSGSS-
GIFGAIAGFIENGWEGMVDGWYGFRHQNSEGRGQAADLKSTQAAIDQINGKLNRLIGKTNEKFH
QIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEKTKKQLRE
NAEDMGNGCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWISFA
ISCFLLCVALLGFIMWACQKGNIRCNICI

Fig. 22

Dãy số	Mẫu	Mô tả
1	496-59-001A	H5/Indo CTRL (# 489)
2	496-59-001B	H5/Indo CTRL (# 489)
3	496-59-002A	H5/Indo PRL- (# 928)
4	496-59-002B	H5/Indo PRL- (# 928)
5	496-59-003A	H5/Indo TETR (# 676)
6	496-59-003B	H5/Indo TETR (# 676)
7	496-59-004A	H5/Indo TETQ (# 766)
8	496-59-004B	H5/Indo TETQ (# 766)
9	MW	
10	C+	H5 25 ng

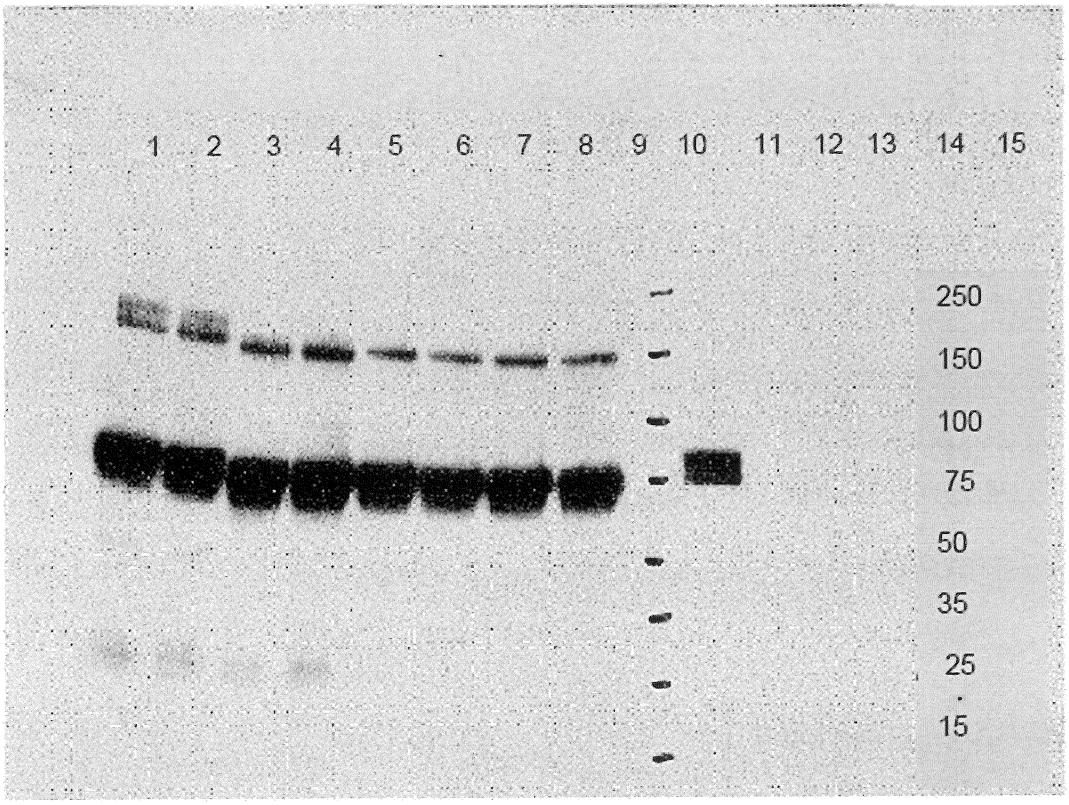


Fig.23: A-2X35S/CPMV-HT/H5 từ A/Indonesia/5/2005 với đột biến vị trí tách TETR (Cấu trúc số 676)

Fig.23A: SEQ ID NO: 74 (MutCleavage-H5(Indo).r

TAGTCCTCTTGTCTCTGTTTGAGGGCTATTTCTGAGCCCTGTTGC

Fig.23B: SEQ ID NO : 75 MutCleavage-H5(Indo).c

TAGCCCTCAAACAGAGACAAGAGGACTATTTGGAGCTATAGCAGG

Fig.23C: SEQ ID NO : 76

Cát xét biểu hiện số 676 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. H5 từ cúm A/Indonesia/5/2005 (H5N1) thể đột biến vị trí tách TETR được gạch dưới

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTC
AGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCG
GATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGC
TCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGA
CAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTC
CAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTT
GTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTT
TCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACT
TTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAA
GGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACC
CACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATT
GATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGAC
CCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTT
TTGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATC
TCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAG
ATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTT
TGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGT
GGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTA
CGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTT
GCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTC
TTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTA
TATTCTGCCCAAATTTGTCGGGCCCATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCA
GTCTTGTTAAAAGTGATCAGATTTGCATTGGTTACCATGCAAACAATTCAACAGAGCAG

GTTGACACAATCATGGAAAAGAACGTTACTGTTACACATGCCCAAGACATACTGGAAAA
GACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTTAAGAGATT
GTAGTGTAGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCG
GAATGGTCTTACATAGTGGAGAAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGAG
TTTCAACGACTATGAAGAACTGAAACACCTATTGAGCAGAATAAACCATTTTGAGAAAA
TTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCTCATCAGGAGTTAGCTCA
GCATGTCCATACCTGGGAAGTCCCTCCTTTTTTAGAAATGTGGTATGGCTTATCAAAAA
GAACAGTACATACCCAACAATAAAGAAAAGCTACAATAATACCAACCAAGAGGATCTTT
TGGTACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGCTATATCAA
AACCCAACCACCTATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAA
AATAGCTACTAGATCCAAAGTAAACGGGCAAAGTGGAAGGATGGAGTTCTTCTGGACAA
TTTTAAAACCTAATGATGCAATCAACTTCGAGAGTAATGGAAATTTCATTGCTCCAGAA
TATGCATACAAAATTGTCAAGAAAGGGGACTCAGCAATTATGAAAAGTGAATTGGAATA
TGGTAACTGCAACACCAAGTGTCAAACTCCAATGGGGGCGATAAACTCTAGTATGCCAT
TCCACAACATACACCCTCTCACCATCGGGGAATGCCCCAAATATGTGAAATCAAACAGA
TTAGTCCTTGCAACAGGGCTCAGAAATAGCCCTCAAACAGAGACAAGAGGACTATTTGG
AGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGT
ACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTCAAAAG
GCAATAGATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTT
TGAGGCCGTTGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGAGAATTTAAACAAGA
AGATGGAAGACGGGTTTCTAGATGTCTGGACTTATAATGCCGAACTTCTGGTTCTCATG
GAAAATGAGAGAACTCTAGACTTTCATGACTCAAATGTTAAGAACCTCTACGACAAGGT
CCGACTACAGCTTAGGGATAATGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATC
ACAAATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAACTATCCGCAG
TATTCAGAAGAAGCAAGATTAAAAAGAGAGGAAATAAGTGGGGTAAAATTGGAATCAAT
AGGAACTTACCAAATACTGTCAATTTATTCAACAGTGGCGAGTTCCCTAGCACTGGCAA
TCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAGAATT
TGCATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTA
TGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTT
TGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTT
ATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGAT
CGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGAT
GATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCA
TGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATAC
GCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATC
TATGTTACTAGAT

Fig.23D: SEQ ID NO:77

Trình tự axit amin của H5 từ cúm A/Indonesia/5/2005 (H5N1) thể đột biến vị trí tách TETR

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCD
LDGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGSFNDYEELK
HLLSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIK
KSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVN
GQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTKCQ
TPMGAINSSMPFHNIHPLTIGECPKYVKSNRLVLATGLRNSPQTETRGLFGAIAGFIEG
GWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFN
NLERRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDNA
KELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEISGVKLESIGTYQILSI
YSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

Fig.23F

Sự biểu hiện theo sơ đồ của cấu trúc số 676.

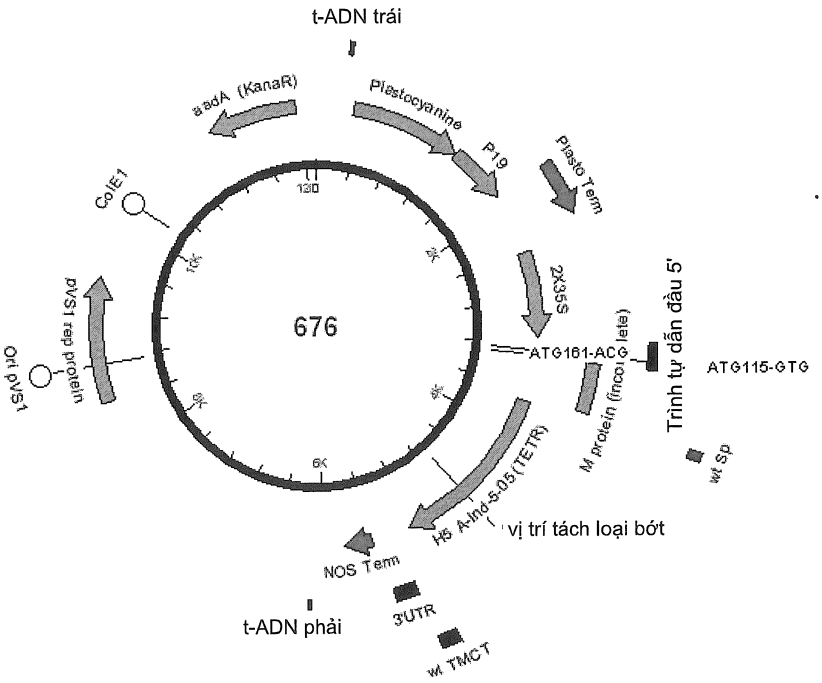


Fig. 24: B-2X35S/CPMV-HT/H5 từ A/Indonesia/5/2005 với đột biến vị trí tách TETR (Cấu trúc số 766)

Fig.24A: SEQ ID NO : 78 H5I505_TETQ.r

TCCAAATAGTCCTTGTGTCTCTGTTTGAGGGCTATTTCTGAGCCCTGT

Fig.24B: SEQ ID NO : 79 H5I505_TETQ.c

AAATAGCCCTCAAACAGAGACACAAGGACTATTTGGAGCTATAGCAGG

Fig.24C: SEQ ID NO : 80

Cát xét biểu hiện số 766 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. H5 từ cúm A/Indonesia/5/2005 (H5N1) thể đột biến vị trí tách được gạch dưới.

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTC
AGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCG
GATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGC
TCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGA
CAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTC
CAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTT
GTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTT
TCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACT
TTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAA
GGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACC
CACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATT
GATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGAC
CCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTT
TTGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATC
TCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAG
ATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTT
TGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGT
GGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTA
CGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTT
GCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTC
TTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTA
TATTCTGCCCAAATTTGTCGGGCCCATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCA
GTCTTGTTAAAAGTGATCAGATTTGCATTGGTTACCATGCAAACAATTCAACAGAGCAG

GTTGACACAATCATGGAAAAGAACGTTACTGTTACACATGCCCAAGACATACTGGAAAA
GACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTTAAGAGATT
GTAGTGTAGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCG
GAATGGTCTTACATAGTGGAGAAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGAG
TTTCAACGACTATGAAGAACTGAAACACCTATTGAGCAGAATAAACCATTTTGAGAAAA
TTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCTCATCAGGAGTTAGCTCA
GCATGTCCATACCTGGGAAGTCCCTCCTTTTTTAGAAATGTGGTATGGCTTATCAAAAA
GAACAGTACATACCCAACAATAAAGAAAAGCTACAATAATACCAACCAAGAGGATCTTT
TGGTACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGCTATATCAA
AACCCAACCACCTATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAA
AATAGCTACTAGATCCAAAGTAAACGGGCAAAGTGGAAGGATGGAGTTCTTCTGGACAA
TTTTAAAACCTAATGATGCAATCAACTTCGAGAGTAATGGAAATTTCATTGCTCCAGAA
TATGCATACAAAATTGTCAAGAAAGGGGACTCAGCAATTATGAAAAGTGAATTGGAATA
TGGTAACTGCAACACCAAGTGTCAAACTCCAATGGGGGCGATAAACTCTAGTATGCCAT
TCCACAACATACACCCTCTCACCATCGGGGAATGCCCCAAATATGTGAAATCAAACAGA
TTAGTCCTTGCAACAGGGCTCAGAAATAGCCCTCAAACAGAGACACAAGGACTATTTGG
AGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGT
ACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTCAAAAG
GCAATAGATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTT
TGAGGCCGTTGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGAGAATTTAAACAAGA
AGATGGAAGACGGGTTTCTAGATGTCTGGACTTATAATGCCGAACTTCTGGTTCTCATG
GAAAATGAGAGAACTCTAGACTTTCATGACTCAAATGTTAAGAACCTCTACGACAAGGT
CCGACTACAGCTTAGGGATAATGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATC
ACAAATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAACTATCCGCAG
TATTCAGAAGAAGCAAGATTAAAAAGAGAGGAAATAAGTGGGGTAAAATTGGAATCAAT
AGGAACTTACCAAATACTGTCAATTTATTCAACAGTGGCGAGTTCCCTAGCACTGGCAA
TCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAGAATT
TGCATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTA
TGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTT
TGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTT
ATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGAT
CGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGAT
GATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCA
TGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATAC
GCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATC
TATGTTACTAGAT

Fig.24D: SEQ ID NO:81

Trình tự axit amin của H5 từ cúm A/Indonesia/5/2005 (H5N1) thể đột biến vị trí tách TETQ.

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCD
LDGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGSFNDYEELK
HLLSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIK
KSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVN
GQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTKCQ
TPMGAINSSMPFHNIHPLTIGECPKYVKSNRLVLATGLRNSPQTETQGLFGAIAGFIEG
GWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFN
NLERRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDNA
KELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEISGVKLESIGTYQILSI
YSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

Fig.24E

Sự biểu diễn theo sơ đồ của cấu trúc số 766.

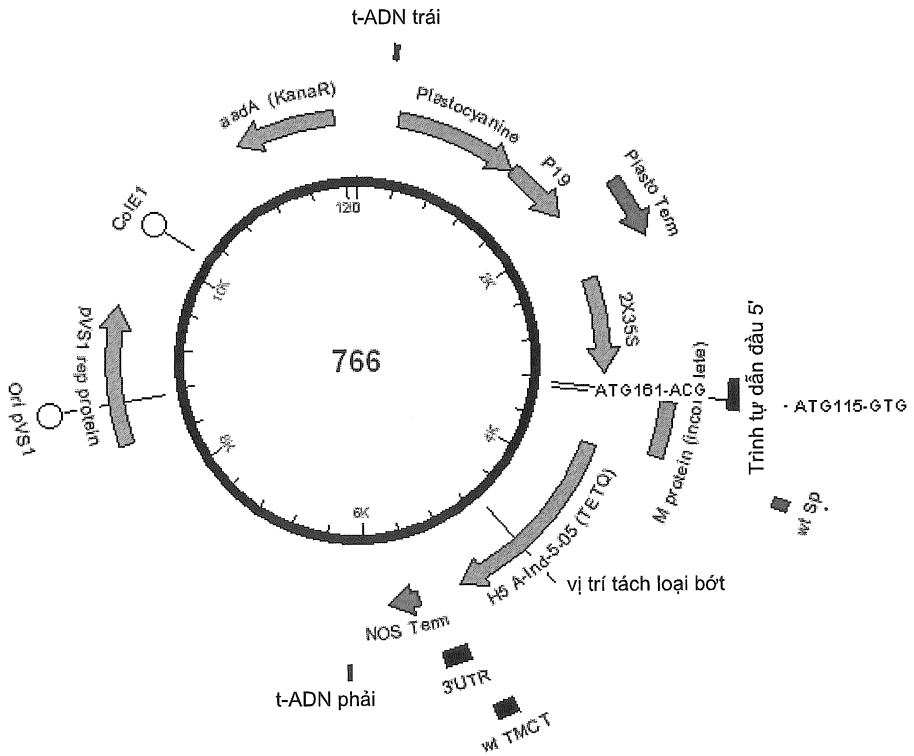


Fig.25: C-2X35S/CPMV-HT/H5 từ A/Indonesia/5/2005 với vòng phân giải protein loại bớt (Cấu trúc số 928)

Fig.25A: SEQ ID NO : 82 H5I505(PrL-).r

CTGCCATCCTCCGCCAGGGCTATTTCTGAGCCCTGTTGCAAGGACTAATC

Fig.25B: SEQ ID NO : 83 H5I505(PrL-).c

GAAATAGCCCTGGCGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGTA

Fig.25C: SEQ ID NO : 84

Cát xét biểu hiện số 928 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. H5 từ cúm A/Indonesia/5/2005 (H5N1) với vòng phân giải protein loại bớt được gạch dưới

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTC
AGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCG
GATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGC
TCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGA
CAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTC
CAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTT
GTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTT
TCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACT
TTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAA
GGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACC
CACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATT
GATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGAC
CCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTT
TTGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATC
TCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAG
ATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTT
TGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGT
GGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTA
CGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTT
GCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTC
TTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTA
TATTCTGCCCAAATTTGTCGGGCCCATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCA
GTCTTGTTAAAAGTGATCAGATTTGCATTGGTTACCATGCAAACAATTCAACAGAGCAG

GTTGACACAATCATGGAAAAGAACGTTACTGTTACACATGCCCAAGACATACTGGAAAA
GACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTTAAGAGATT
GTAGTGTAGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCG
GAATGGTCTTACATAGTGGAGAAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGAG
TTTCAACGACTATGAAGAACTGAAACACCTATTGAGCAGAATAAACCATTTTGAGAAAA
TTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCTCATCAGGAGTTAGCTCA
GCATGTCCATACCTGGGAAGTCCCTCCTTTTTTAGAAATGTGGTATGGCTTATCAAAAA
GAACAGTACATACCCAACAATAAAGAAAAGCTACAATAATACCAACCAAGAGGATCTTT
TGGTACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGCTATATCAA
AACCCAACCACCTATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAA
AATAGCTACTAGATCCAAAGTAAACGGGCAAAGTGGAAGGATGGAGTTCTTCTGGACAA
TTTTAAAACCTAATGATGCAATCAACTTCGAGAGTAATGGAAATTTCATTGCTCCAGAA
TATGCATACAAAATTGTCAAGAAAGGGGACTCAGCAATTATGAAAAGTGAATTGGAATA
TGGTAACTGCAACACCAAGTGTCAAACTCCAATGGGGGCGATAAACTCTAGTATGCCAT
TCCACAACATACACCCTCTCACCATCGGGGAATGCCCCAAATATGTGAAATCAAACAGA
TTAGTCCTTGCAACAGGGCTCAGAAATAGCCCTGGCGGAGGATGGCAGGGAATGGTAGA
TGGTTGGTATGGGTACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAG
AATCCACTCAAAAGGCAATAGATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAA
ATGAACACTCAGTTTGAGGCCGTTGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGA
GAATTTAAACAAGAAGATGGAAGACGGGTTTCTAGATGTCTGGACTTATAATGCCGAAC
TTCTGGTTCTCATGGAAAATGAGAGAACTCTAGACTTTCATGACTCAAATGTTAAGAAC
CTCTACGACAAGGTCCGACTACAGCTTAGGGATAATGCAAAGGAGCTGGGTAACGGTTG
TTTCGAGTTCTATCACAAATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGT
ACAACTATCCGCAGTATTCAGAAGAAGCAAGATTAAAAAGAGAGGAAATAAGTGGGGTA
AAATTGGAATCAATAGGAACTTACCAAATACTGTCAATTTATTCAACAGTGGCGAGTTC
CCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGT
TACAATGCAGAATTTGCATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTC
GGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATG
TAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAA
AAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTT
ATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTT
GCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAAT
TAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAAT
TATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCG
CGCGCGGTGTCATCTATGTTACTAGAT

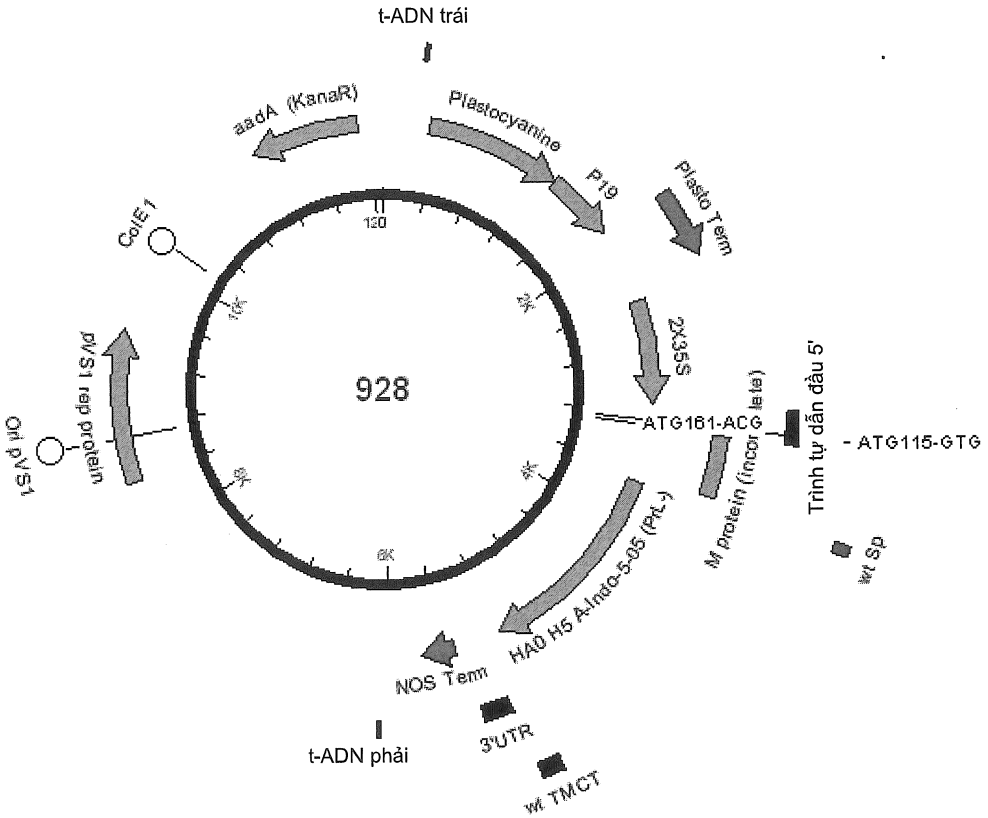
Fig.25D: SEQ ID NO:85

Trình tự axit amin của H5 từ cúm A/Indonesia/5/2005 (H5N1) với vòng phân giải protein loại bớt

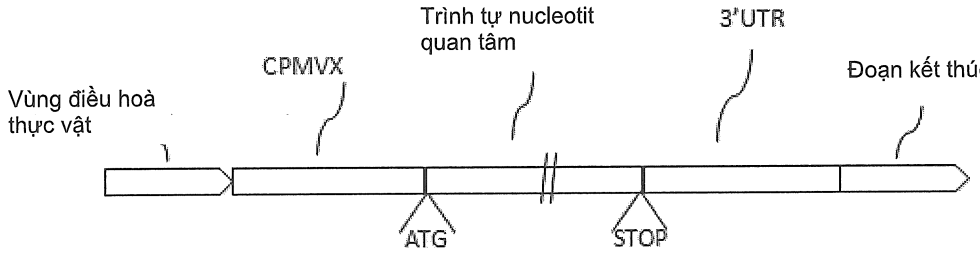
MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCD
LDGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPTNDLCYPGSFNDYEELK
HLLSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIK
KSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIATRSKVN
GQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTKCQ
TPMGAINSSMPFHNIHPLTIGECPKYVKSNRLVLATGLRNSPGGGWQGMVDGWYGYHHS
NEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLNKKMED
GFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDNAKELGNGCFEFYHKCD
NECMESIRNGTYNYPQYSEEARLKREEISGVKLESIGTYQILSIYSTVASSLALAIMMA
GLSLWMCSNGSLQCRICI

Fig.25E

Sự biểu diễn theo sơ đồ của cấu trúc số 928.



Cấu trúc bao gồm CPMV160



Cấu trúc bao gồm CPMV160+

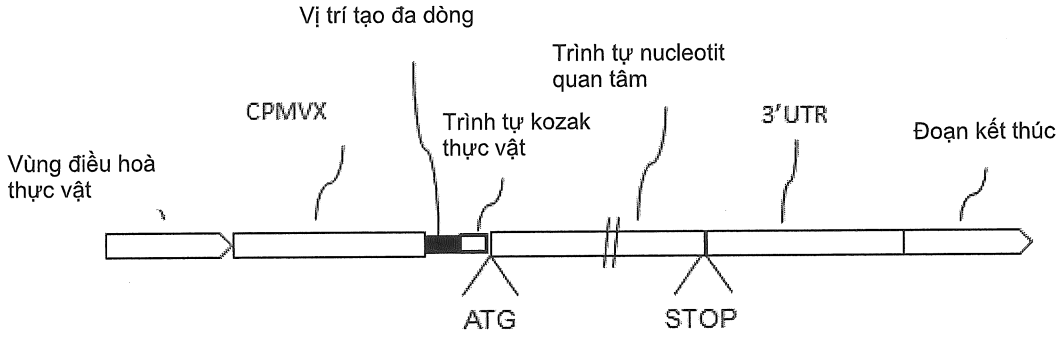


Fig.26A

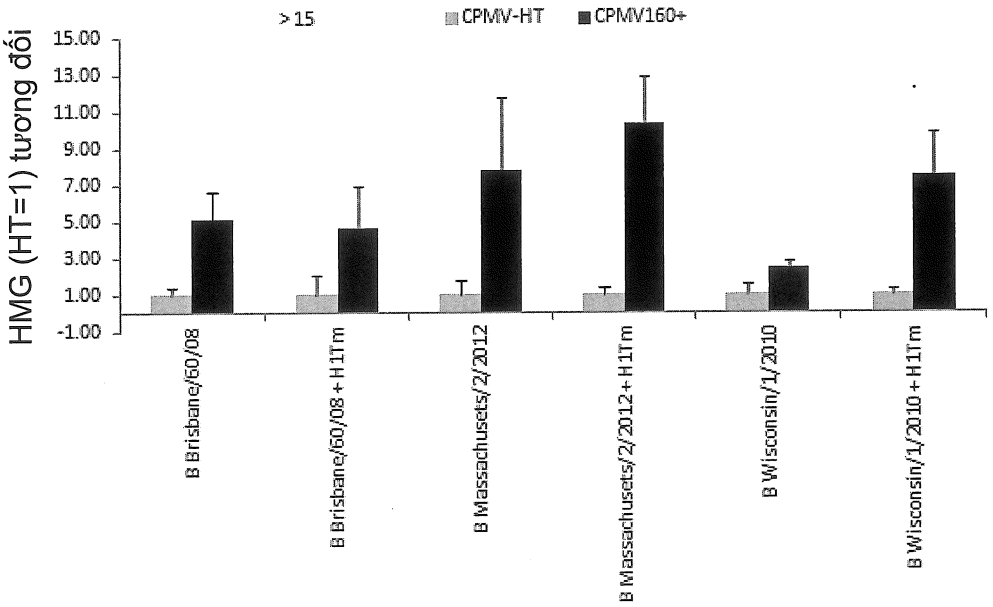


Fig.26B

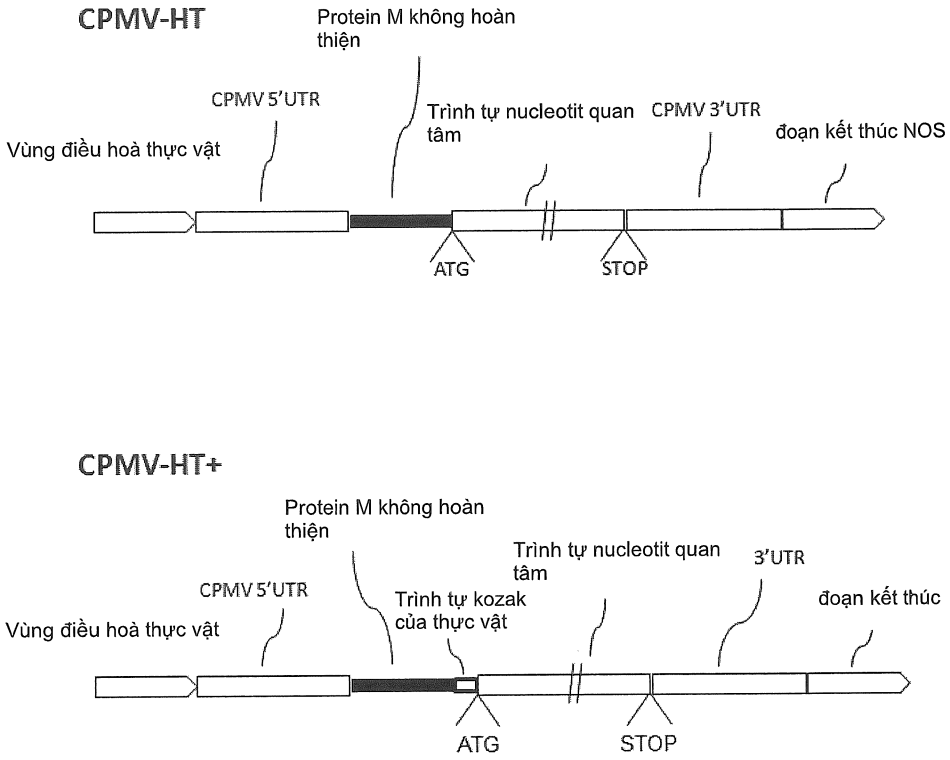


Fig.27A

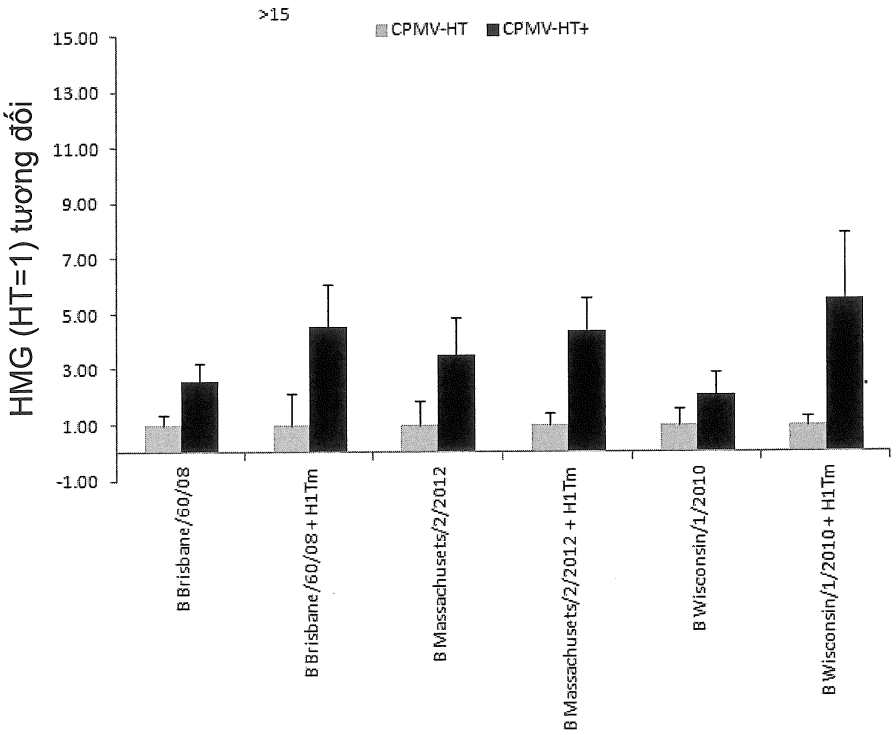
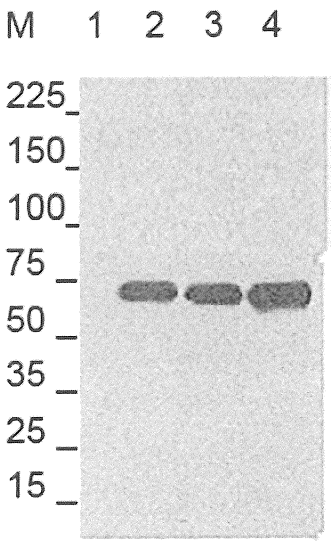


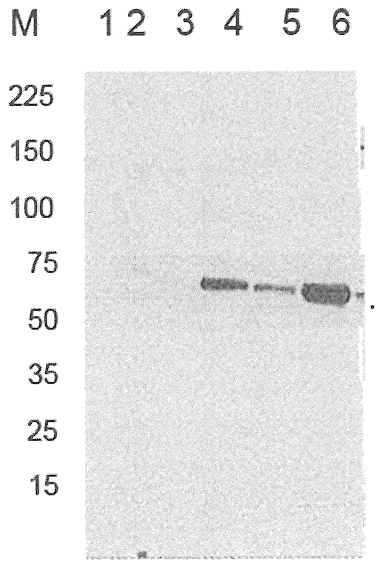
Fig.27B

FIG 28A



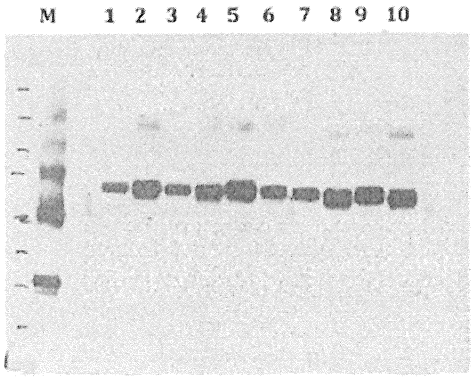
1: wT HT+
2: wT 160+
3: PrL HT+
4: PrL 160+

28B



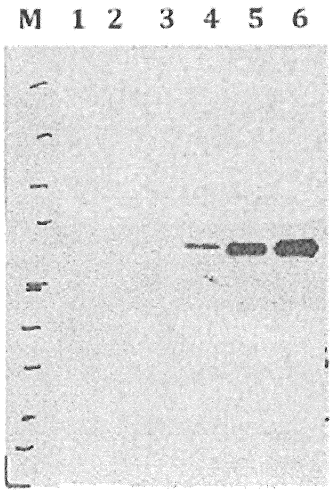
3: PrL- 160+
4: PrL- B/160+
5: PrL-+H/c 160+
6: PrL-+H/c B/160+

FIG 28C



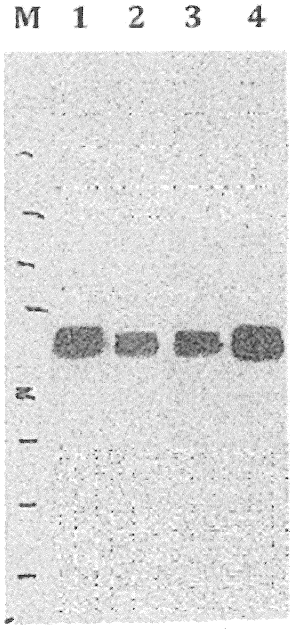
1: wT HT
2: wT B/HT
3: wT HT+
4: wT B/HT
5: wT 160+
6: wT B/160+
7: PrL- HT+
8: PrL- B/HT+
9: PrL- 160+
10: PrL- B/160+

FIG 28D



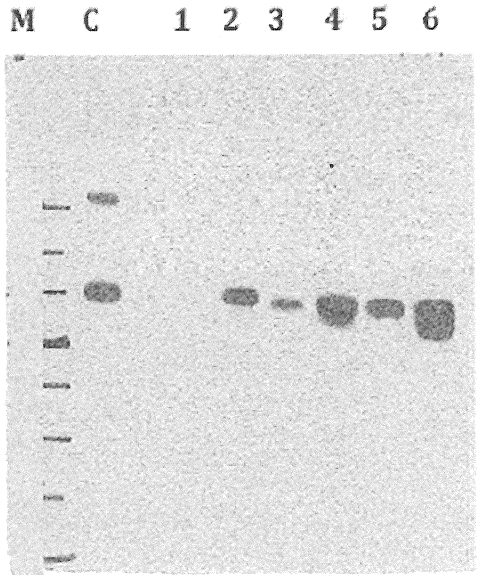
1: wT HT
2: wT HT+
3: wT 160+
4: PrL HT
5: PrL HT+
6: PrL 160+

FIG 28E



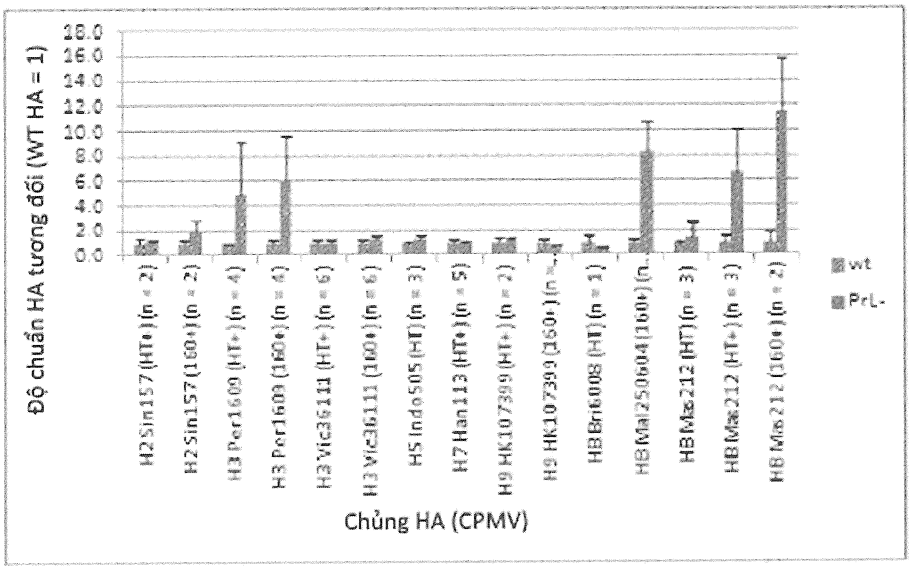
1: wT 160+
2: wT HT+
3: PrL HT+
4: PrL 160+

28F



1: HT/wT
2: B/HT wT
3: HT+/PrL-
4: B/HT+/PrL-
5. HT+/PrL +H/C
6. B/HT+/PrL+H/C

FIG. 29A



F-2X35S/CPMV-HT/PDISP/HA B Brisbane/NOS (Cấu trúc số 1029)

Fig.30A

IF-S2+S4-B Bris.c (SEQ ID NO: 86)

TCTCAGATCTTCGCCGATCGAATCTGCACTGGAATAACAT

Fig.30B

IF-S1a4-B Bris.r (SEQ ID NO: 87)

ACTAAAGAAAATAGGCCTTTATAGACAGATGGAGCAAGAAACA

Fig.30C

Gen HA B Brisbane được tổng hợp (tương ứng với nt 34-1791 từ số truy cập Genbank FJ766840) (SEQ ID NO:88)

ATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGAA
TAACATCGTCAAACTCACCACATGTCGTCAAAACTGCTACTCAAGGGGAGGTCAATGTGACTGG
TGTAATACCACTGACAACAACACCCACCAAATCTCATTTTGCAAATCTCAAAGGAACAGAAACC
AGGGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATCTGGACGTAGCCTTGGGCAGACCAA
AATGCACGGGGAAAATACCCTCGGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTACATC
TGGGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCCTAACCTTCTCCGAGGA
TACGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAAAATGCACCAGGAGGACCCT
ACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTTCGCAACAAT
GGCTTGGGCCGTCCCAAAAAACGACAAAAACAAAACAGCAACAAATCCATTAACAATAGAAGTA
CCATACATTTGTACAGAAGGAGAAGACCAAATTACCGTTTGGGGGTTCCACTCTGACAACGAGA
CCCAAATGGCAAAGCTCTATGGGGACTCAAAGCCCCAGAAGTTCACCTCATCTGCCAACGGAGT
GACCACACATTACGTTTCACAGATTGGTGGCTTCCCAAATCAAACAGAAGACGGAGGACTACCA
CAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGGGAAAACAGGAACAATTACCT
ATCAAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTGCGCAAGTGGCAGGAGCAAGGTAATAAA
AGGATCCTTGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTGGATTAAACAAA
AGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAATATGGGTGAAAA
CACCCTTGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGCAAAACTATTAAAGGAAAGGGG
TTTCTTCGGAGCTATTGCTGGTTTCTTAGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCAC
GGATACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGG
CCATAAACAAGATAACAAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAG
ACTAAGCGGTGCCATGGATGAACTCCACAACGAAATACTAGAACTAGATGAGAAAGTGGATGAT
CTCAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAGTCCTGCTTTCCAATGAAGGAATAA
TAAACAGTGAAGATGAACATCTCTTGGCGCTTGAAAGAAAGCTGAAGAAAATGCTGGGCCCCTC
TGCTGTAGAGATAGGGAATGGATGCTTTGAAACCAAACACAAGTGCAACCAGACCTGTCTCGAC
AGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTCTCCCCACCTTTGATTCACTGAATA

TTACTGCTGCATCTTTAAATGACGATGGATTGGATAATCATACTATACTGCTTTACTACTCAAC
TGCTGCCTCCAGTTTGGCTGTAACACTGATGATAGCTATCTTTGTTGTTTATATGGTCTCCAGA
GACAATGTTTCTTGCTCCATCTGTCTATAA

Fig.30D

Cát xét biểu hiện số 1029 từ đoạn khởi đầu 2x35S đến đoạn kết thúc NOS. PDISP/HA từ cúm B/Brisbane/60/2008 được gạch dưới. (SEQ ID NO:89)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAG
ACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCAT
CATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGAC
CCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGA
TTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTC
TCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGAT
TCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAA
ATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAA
GATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGC
AAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCA
AGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTT
TGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTT
AAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCG
GCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGC
GGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTT
GCTGGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGC
AATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGA
TTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGG
AGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGGCGAAAAACGTTG
CGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCGAAT
CTGCACTGGAATAACATCGTCAAACTCACCACATGTCGTCAAAACTGCTACTCAAGGGGAGGTC
AATGTGACTGGTGTAATACCACTGACAACAACACCCACCAAATCTCATTTTGCAAATCTCAAAG
GAACAGAAACCAGGGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATCTGGACGTAGCCTT
GGGCAGACCAAAATGCACGGGGAAAATACCCTCGGCAAGAGTTTCAATACTCCATGAAGTCAGA
CCTGTTACATCTGGGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCCTAACC
TTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAAAATGCACC
AGGAGGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTT
TTCGCAACAATGGCTTGGGCCGTCCCAAAAAACGACAAAAACAAAACAGCAACAAATCCATTAA
CAATAGAAGTACCATACATTTGTACAGAAGGAGAAGACCAAATTACCGTTTGGGGGTTCCACTC
TGACAACGAGACCCAAATGGCAAAGCTCTATGGGGACTCAAAGCCCCAGAAGTTCACCTCATCT
GCCAACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCAAATCAAACAGAAGACG
GAGGACTACCACAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGGGAAAACAGG
AACAATTACCTATCAAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTGCGCAAGTGGCAGGAGC
AAGGTAATAAAAGGATCCTTGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTG
GATTAAACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAAT
ATGGGTGAAAACACCCTTGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGCAAAACTATTA
AAGGAAAGGGGTTTCTTCGGAGCTATTGCTGGTTTCTTAGAAGGAGGATGGGAAGGAATGATTG

CAGGTTGGCACGGATACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAG
CACTCAAGAGGCCATAAACAAGATAACAAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAG
AATCTTCAAAGACTAAGCGGTGCCATGGATGAACTCCACAACGAAATACTAGAACTAGATGAGA
AAGTGGATGATCTCAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAGTCCTGCTTTCCAA
TGAAGGAATAATAAACAGTGAAGATGAACATCTCTTGGCGCTTGAAAGAAAGCTGAAGAAAATG
CTGGGCCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTGAAACCAAACACAAGTGCAACCAGA
CCTGTCTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTCTCCCCACCTTTGA
TTCACTGAATATTACTGCTGCATCTTTAAATGACGATGGATTGGATAATCATACTATACTGCTT
TACTACTCAACTGCTGCCTCCAGTTTGGCTGTAACACTGATGATAGCTATCTTTGTTGTTTATA
TGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGTCTATAAAGGCCTATTTTCTTTAGTTTGAA
TTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTT
ATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACAC
AAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTAT
CGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTC
TTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATG
CATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCG
ATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTAC
TAGAT

Fig.30E:

Trình tự axit amin của PDISP/HA từ cúm
B/Brisbane/60/2008 SEQ ID NO: 90

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSH
FANLKGTETRGKLCPKCLNCTDLDVALGRPKCTGKIPSARVSILHEVRPVTSGCFPIMHDRTKI
RQLPNLLRGYEHIRLSTHNVINAENAPGGPYKIGTSGSCPNITNGNGFFATMAWAVPKNDKNKT
ATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFP
NQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCL
HEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLKERGFFGAIAGFLEGG
WEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEI
LELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVEIGNGCFETK
HKCNQTCLDRIAAGTFDAGEFSLPTFDSLNITAASLNDDGLDNHTILLYYSTAASSLAVTLMIA
IFVVYMVSRDNVSCSICL

Fig. 30F

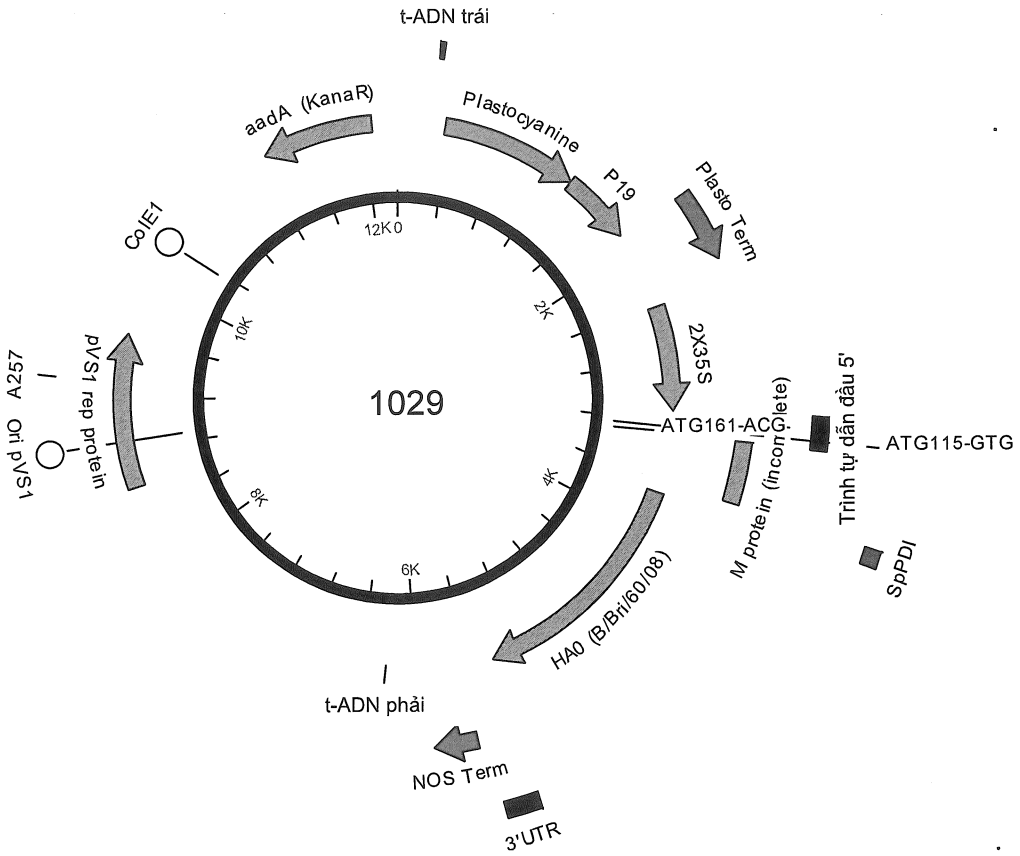

t-ADN trái
aadA (KanaR)
Plastocyanine
P19
Plasto Term
2X35S
ATG161-ACG
ATG115-GTG
Trình tự dẫn đầu 5'
M protein (incomplete)
SpPDI
HA0 (B/Bri/60/08)
t-ADN phải
NOS Term
3'UTR
CoIE1
pVS1 rep protein
Ori pVS1
A257
1029
12K0
2K
4K
6K
8K
10K

Fig.31
K-2X35S/CPMV HT (cấu trúc số 1039) và HT* (cấu trúc số 1829) đối với PDISP/HA B Brisbane (PrL-)

Fig.31A: SEQ ID NO:91 Trình tự nucleotit của PDISP/HA B Brisbane (PrL-).

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGA
TCTTCGCCGATCGAATCTGCACTGGAATAACATCGTCAAACTCACCACATGTCGTCAAAACTGC
TACTCAAGGGGAGGTCAATGTGACTGGTGTAATACCACTGACAACAACACCCACCAAATCTCAT
TTTGCAAATCTCAAAGGAACAGAAACCAGGGGGAAACTATGCCCAAAATGCCTCAACTGCACAG
ATCTGGACGTAGCCTTGGGCAGACCAAAATGCACGGGGAAAATACCCTCGGCAAGAGTTTCAAT
ACTCCATGAAGTCAGACCTGTTACATCTGGGTGCTTTCCTATAATGCACGACAGAACAAAAATT
AGACAGCTGCCTAACCTTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTATCA
ATGCAGAAAATGCACCAGGAGGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTAC
CAATGGAAACGGATTTTTCGCAACAATGGCTTGGGCCGTCCCAAAAAACGACAAAAACAAAACA
GCAACAAATCCATTAACAATAGAAGTACCATACATTTGTACAGAAGGAGAAGACCAAATTACCG
TTTGGGGGTTCCACTCTGACAACGAGACCCAAATGGCAAAGCTCTATGGGGACTCAAAGCCCCA
GAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCA
AATCAAACAGAAGACGGAGGACTACCACAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAA
AATCTGGGAAAACAGGAACAATTACCTATCAAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTG
CGCAAGTGGCAGGAGCAAGGTAATAAAAGGATCCTTGCCTTTAATTGGAGAAGCAGATTGCCTC
CACGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCA
TAGGAAATTGCCCAATATGGGTGAAAACACCCTTGAAGCTGGCCAATGGAACCAAATATAGACC
TCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCCCATGGGGCACAT
GGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAACAAGATAACAAAAAATC
TCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCCATGGATGAACT
CCACAACGAAATACTAGAACTAGATGAGAAAGTGGATGATCTCAGAGCTGATACAATAAGCTCA
CAAATAGAACTCGCAGTCCTGCTTTCCAATGAAGGAATAATAAACAGTGAAGATGAACATCTCT
TGGCGCTTGAAAGAAAGCTGAAGAAAATGCTGGGCCCCTCTGCTGTAGAGATAGGGAATGGATG
CTTTGAAACCAAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGGTACCTTTGAT
GCAGGAGAATTTTCTCTCCCCACCTTTGATTCACTGAATATTACTGCTGCATCTTTAAATGACG
ATGGATTGGATAATCATACTATACTGCTTTACTACTCAACTGCTGCCTCCAGTTTGGCTGTAAC
ACTGATGATAGCTATCTTTGTTGTTTATATGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGT
CTATAA

Fig.31B: SEQ ID NO:92

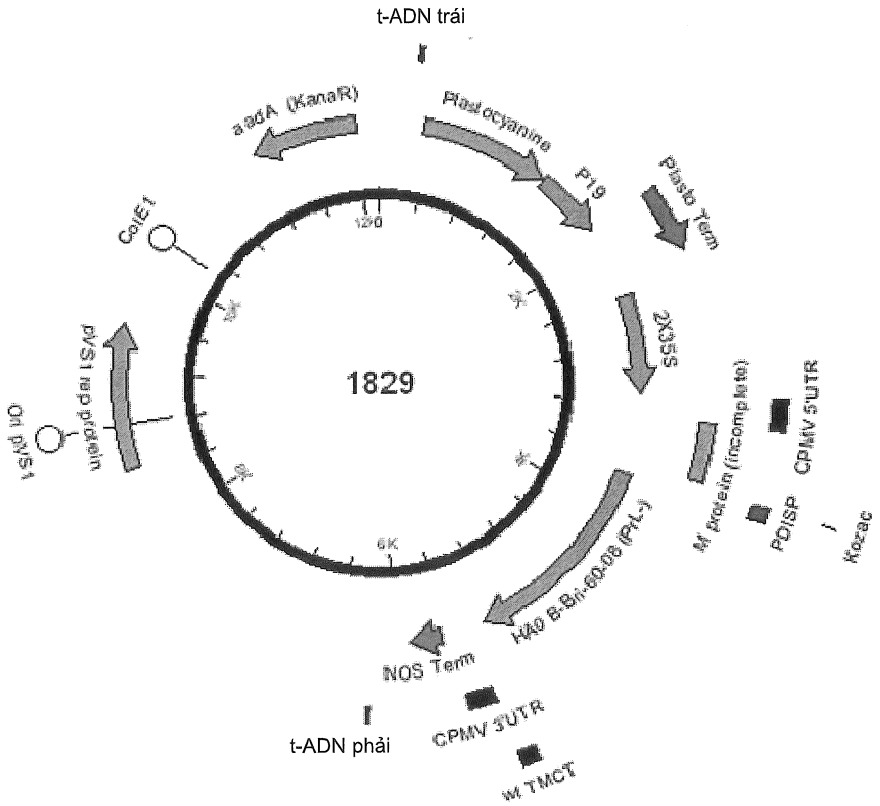
Trình tự axit amin của PDISP/HA B Brisbane (PrL-).

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSH
FANLKGTETRGKLCPKCLNCTDLDVALGRPKCTGKIPSARVSILHEVRPVTSGCFPIMHDRTKI
RQLPNLLRGYEHIRLSTHNVINAENAPGGPYKIGTSGSCPNITNGNGFFATMAWAVPKNDKNKT
ATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFP
NQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCL

HEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAH
GVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISS
QIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFD
AGEFSLPTFDSLNITAASLNDDGLDNHTILLYYSTAASSLAVTLMIAIFVVYMVSRDNVSCSIC
L*

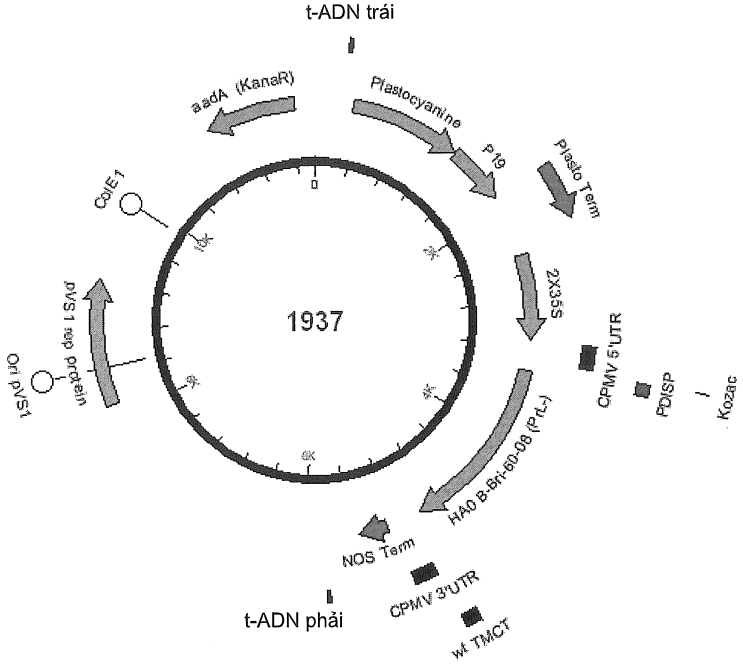
Fig.31C

Sự biểu diễn theo sơ đồ của cấu trúc số 1829 (2X35S/CPMV HT*)



2X35S/CPMV HT (Cấu trúc số 1039) và HT*(-Mprot) (cấu trúc số 1937) đối với PDISP/HA B Brisbane (PrL-)

Fig.32A
Sự biểu diên theo sơ đồ của cấu trúc số 1937 (2X35S/CPMV HT*(-Mprot))



2X35S/CPMV HT (cấu trúc số 1067) và HT*(-Mprot) (cấu trúc số 1977) đối với PDISP/HA B Brisbane (PrL-)+H1 California TMCT

Fig.33A: SEQ ID NO:95
Trình tự nucleotit của DPISP/HA B Brisbane (PrL-)+H1 California TMCT

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCGA
ATCTGCACTGGAATAACATCGTCAAACTCACCACATGTCGTCAAAACTGCTACTCAAGGGGAGGTCAATGTGACTG
GTGTAATACCACTGACAACAACACCCACCAAATCTCATTTTGCAAATCTCAAAGGAACAGAAACCAGGGGGAAAC
TATGCCCAAAATGCCTCAACTGCACAGATCTGGACGTAGCCTTGGGCAGACCAAAATGCACGGGGAAAATACCCT
CGGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTACATCTGGGTGCTTTCCTATAATGCACGACAGAACAAA
AATTAGACAGCTGCCTAACCTTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAA
AATGCACCAGGAGGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTTCG
CAACAATGGCTTGGGCCGTCCCAAAAAACGACAAAAACAAAACAGCAACAAATCCATTAACAATAGAAGTACCA
TACATTTGTACAGAAGGAGAAGACCAAATTACCGTTTGGGGGTTCCACTCTGACAACGAGACCCAAATGGCAAAG
CTCTATGGGGACTCAAAGCCCCAGAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGTTTCACAGATTG
GTGGCTTCCCAAATCAAACAGAAGACGGAGGACTACCACAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAA
AATCTGGGAAAACAGGAACAATTACCTATCAAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTGCGCAAGTGGCA
GGAGCAAGGTAATAAAAGGATCCTTGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTGGATTAA
ACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAATATGGGTGAAAACACCCT
TGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGAT
ACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAACAAGATA
ACAAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCCATGGATGAACTC
CACAACGAAATACTAGAACTAGATGAGAAAGTGGATGATCTCAGAGCTGATACAATAAGCTCACAAATAGAACTC
GCAGTCCTGCTTTCCAATGAAGGAATAATAAACAGTGAAGATGAACATCTCTTGGCGCTTGAAAGAAAGCTGAAG
AAAATGCTGGGCCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTGAAACCAAACACAAGTGCAACCAGACCTGT
CTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTCTCCCCACCTTTGATTCACTGAATATTACTGC
TGCATCTTTAAATGACGATGGATTGGATAATTACCAGATTTTGGCGATCTATTCAACTGTCGCCAGTTCATTGGTAC
TGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAA

Fig.33B: SEQ ID NO : 96
Trình tự axit amin của PDISP/HA B Brisbane (PrL-)+H1 California TMCT.

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKGTETRGKLCPKC
LNCTDLDVALGRPKCTGKIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAENAPGGPYKIG
TSGSCPNITNGNGFFATMAWAVPKNDKNKTATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSA
NGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCL
HEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQ
EAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKK
MLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSLNITAASLNDDGLDNYQILAIYSTVASSLVLVVSL
GAISFWMCSNGSLQCRICI*

Fig.33C
Sự biểu diễn theo sơ đồ của cấu trúc số 1067 (2X35S/CPMV HT)

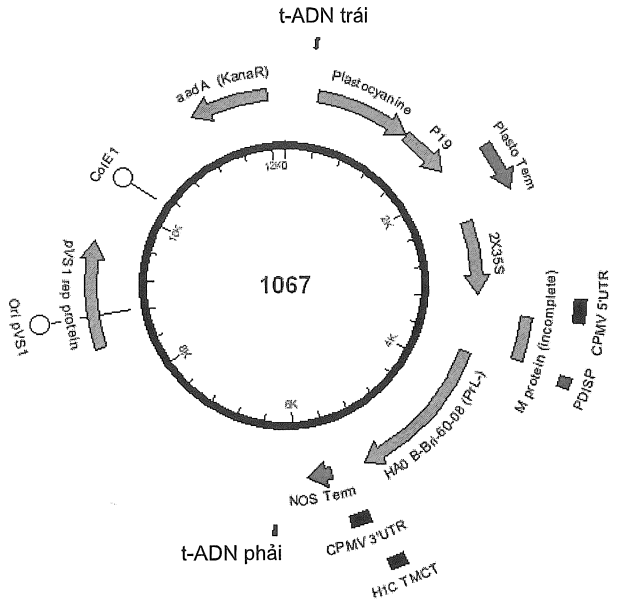
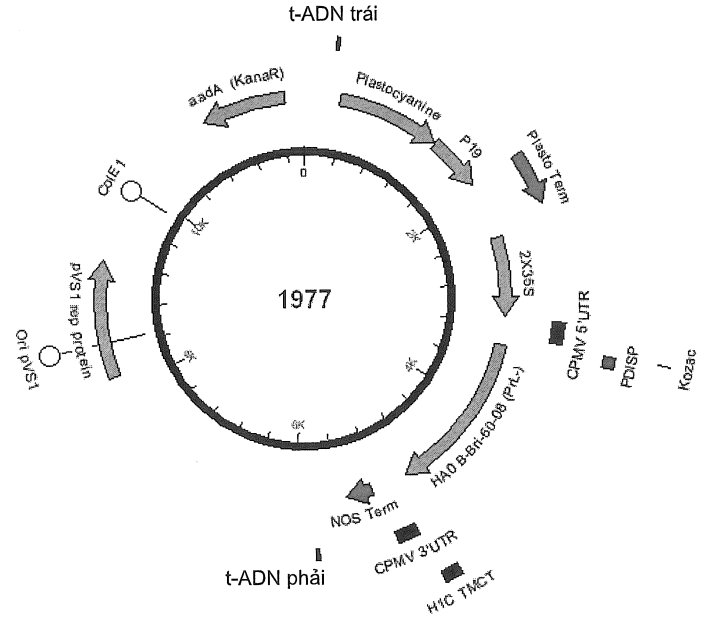


Fig.33D
Sự biểu diễn theo sơ đồ của cấu trúc số 1977 (2X35S/CPMV HT*(-Mprot))



2X35S/CPMV HT (cấu trúc số 2072) và HT*(-Mprot) (cấu trúc số 2050) đối với PDISP/HA B Massachussetts (PrL-)

Fig.34A: SEO ID NO : 97

Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-).

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCG
AATCTGCACTGGGATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTG
GTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAAGACCAGAGGGAAACTA
TGCCCAGACTGTCTCAACTGTACAGATCTGGATGTGGCCCTGGGCAGGCCAATGTGTGTGGGAACTACACCTTCTGC
GAAAGCTTCAATACTTCACGAAGTCAGACCTGTTACATCCGGGTGCTTCCCTATAATGCACGACAGAACAAAAATCA
GGCAACTAGCCAATCTTCTCAGAGGATATGAAAATATCAGGTTATCAACCCAAAACGTTATCGATGCAGAAAAGGCA
CCAGGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAAAGCGGATTTTTCGCAACAAT
GGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCATTAACAGTAGAAGTACCATACATTTGTGCAG
AAGGGGAAGACCAAATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATGGAGACTCA
AATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACATTATGTTTCTCAGATTGGCGGCTTCCCAGATCA
AACAGAAGACGGAGGACTACCACAAAGCGGCAGAATTGTCGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAA
CAATTGTCTATCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGG
TCCTTGCCTTTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACAC
AGGAGAACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTGAAAACACCTTTGAAGCTTGCCAATGGAACCAAAT
ATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCTCACGGAGCACATGGAGTG
GCAGTTGCTGCAGACCTTAAGAGCACACAAGAAGCTATAAACAAGATAACAAAAAATCTCAACTCTTTGAGTGAGCT
AGAAGTAAAGAATCTTCAAAGGCTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAG
TGGATGACCTCAGAGCTGACACTATAAGTTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAAC
AGTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGACATAGGAAA
TGGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAG
AGTTTTCTCTCCCCACTTTTGATTCATTGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAACCATACT
ATACTGCTCTATTACTCAACTGCTGCTTCTAGTTTGGCTGTAACATTGATGCTAGCTATTTTTATTGTTTATATGGT
CTCCAGAGACAACGTTTCATGCTCCATCTGTCTATAA

Fig.34B: . SEO ID NO : 98

Trình tự axit amin của PDISP/HA B Massachussetts (PrL-).

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTKTRGKL
CPDCLNCTDLDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLANLLRGYENIRLSTQNVIDAEKA
PGGPYRLGTSGSCPNATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICAEGEDQITVWGFHSDNKTQMKNLYGDS
NPQKFTSSANGVTTHYVSQIGGFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKG
SLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGV
AVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIIN
SEDEHLLALERKLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNHT
ILLYYSTAASSLAVTLMLAIFIVYMVSRDNVSCSICL*

Fig.34C
Sự biểu diễn theo sơ đồ của cấu trúc số 2072 (2X35S/CPMV HT)

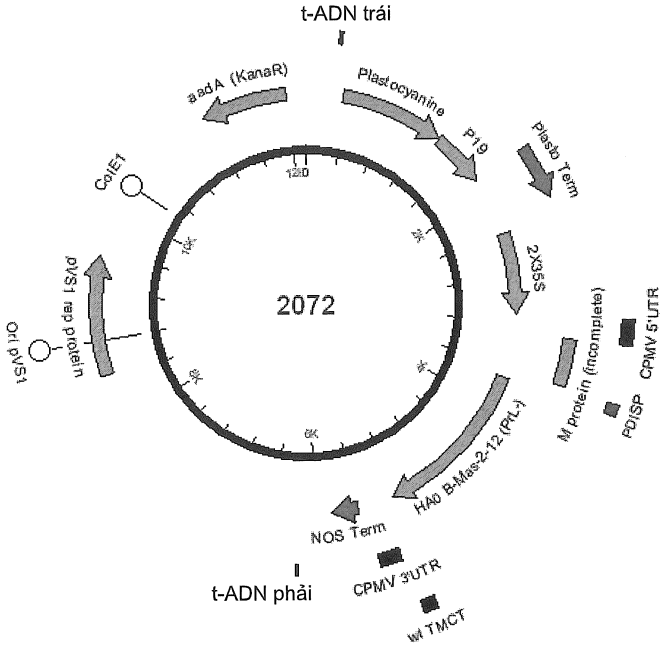
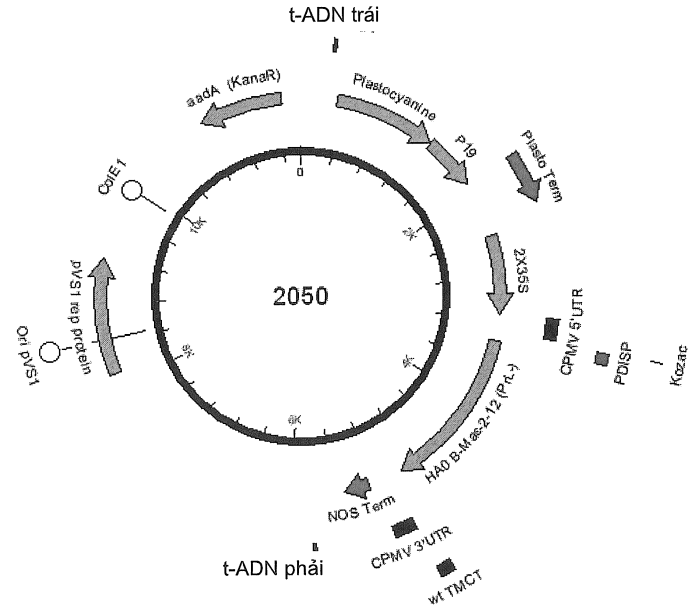


Fig.34D
Sự biểu diễn theo sơ đồ của cấu trúc số 2050 (2X35S/CPMV HT*(-Mprot))



2X35S/CPMV HT (cấu trúc số 2074) và HT*(-Mrpot) (cấu trúc số 2060) đối với PDISP/HA B Massachussetts (PrL-)+H1 California TMCT

Fig.35A: SEQ ID NO : 99

Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT.

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCG
AATCTGCACTGGGATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTG
GTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAAGACCAGAGGGAAACTA
TGCCCAGACTGTCTCAACTGTACAGATCTGGATGTGGCCCTGGGCAGGCCAATGTGTGTGGGAACTACACCTTCTGC
GAAAGCTTCAATACTTCACGAAGTCAGACCTGTTACATCCGGGTGCTTCCCTATAATGCACGACAGAACAAAAATCA
GGCAACTAGCCAATCTTCTCAGAGGATATGAAAATATCAGGTTATCAACCCAAAACGTTATCGATGCAGAAAAGGCA
CCAGGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAAAGCGGATTTTTCGCAACAAT
GGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCATTAACAGTAGAAGTACCATACATTTGTGCAG
AAGGGGAAGACCAAATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATGGAGACTCA
AATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACATTATGTTTCTCAGATTGGCGGCTTCCCAGATCA
AACAGAAGACGGAGGACTACCACAAAGCGGCAGAATTGTCGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAA
CAATTGTCTATCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGG
TCCTTGCCTTTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACAC
AGGAGAACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTGAAAACACCTTTGAAGCTTGCCAATGGAACCAAAT
ATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCTCACGGAGCACATGGAGTG
GCAGTTGCTGCAGACCTTAAGAGCACACAAGAAGCTATAAACAAGATAACAAAAAATCTCAACTCTTTGAGTGAGCT
AGAAGTAAAGAATCTTCAAAGGCTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAG
TGGATGACCTCAGAGCTGACACTATAAGTTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAAC
AGTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGACATAGGAAA
TGGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAG
AGTTTTCTCTCCCCACTTTTGATTCATTGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAACTACCAG
ATTTTGGCGATCTATTCAACTGTCGCCAGTTCATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTG
CTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAA

Fig.35B: SEQ ID NO :100

Trình tự axit amin của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT.

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTKTRGKL
CPDCLNCTDLDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLANLLRGYENIRLSTQNVIDAEKA
PGGPYRLGTSGSCPNATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICAEGEDQITVWGFHSDNKTQMKNLYGDS
NPQKFTSSANGVTTHYVSQIGGFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKG
SLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGV
AVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIIN
SEDEHLLALERKLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNYQ
ILAIYSTVASSLVLVVSLGAISFWMCSNGSLQCRICI*

Fig.35C
Sự biểu diễn theo sơ đồ của cấu trúc số 2074 (2X35S/CPMV HT)

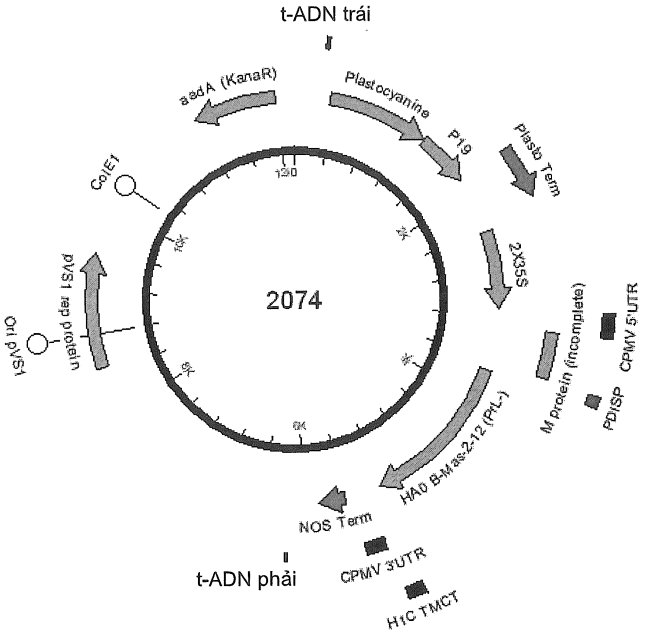
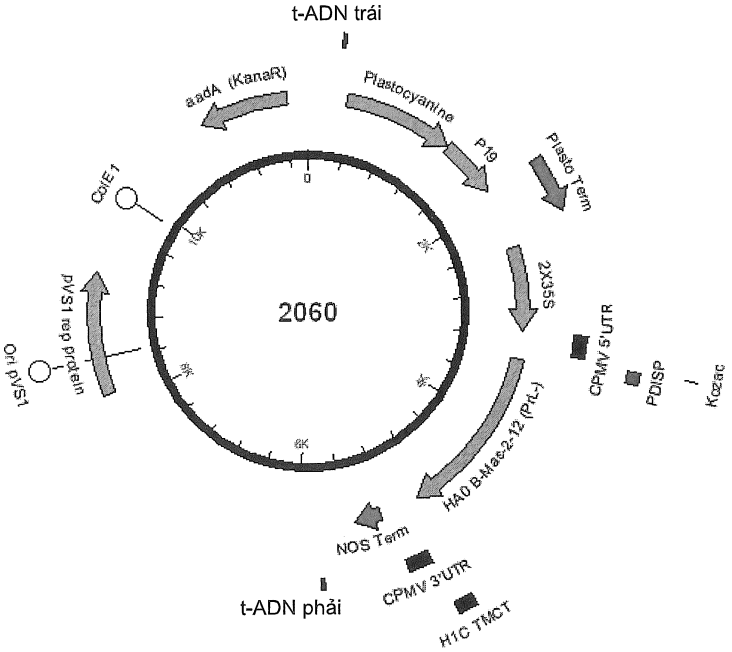


Fig.35D
Sự biểu diễn theo sơ đồ của cấu trúc số 2060 (2X35S/CPMV HT*(-Mprot))



2X35S/CPMV HT (cấu trúc số 1445), HT*(-Mprot) (cấu trúc số 1820) và HT(fl5'UTR) (cấu trúc số 1975) đối với HA B Wisconsin (PrL-)

Fig.36A: SEQ ID NO : 101

Trình tự nucleotit của HA B Wisconsin (PrL-).

ATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGGATAACATCTTCAAA
CTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGCGTGATACCACTGACAACAACACCAA
CAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGCCCGGACTGTCTCAACTGTACAGAT
CTGGATGTGGCCTTGGGCAGGCCAATGTGTGTGGGGACCACACCTTCTGCTAAAGCTTCAATACTCCACGAGGTCAG
ACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAAAATCAGGCAACTACCCAATCTTCTCAGAGGAT
ATGAAAATATCAGGTTATCAACCCAAAACGTTATCGATGCAGAAAAAGCACCAGGAGGACCCTACAGACTTGGAACC
TCAGGATCTTGCCCTAACGCTACCAGTAAAATCGGATTTTTTGCAACAATGGCTTGGGCTGTCCCAAAGGACAACTA
CAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAAATTACTGTTTGGG
GGTTCCATTCAGATAACAAAACCCAAATGAAGAGCCTCTATGGAGACTCAAATCCTCAAAAGTTCACCTCATCTGCT
AATGGAGTAACCACACATTATGTTTCTCAGATTGGCGACTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAG
CGGCAGAATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTATCAAAGAGGTGTTTTGT
TGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCATTGCCTTTAATTGGTGAAGCAGAT
TGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAAA
TTGCCCAATATGGGTAAAAACACCTTTGAAGCTTGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAG
GAATGATTGCAGGTTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCAGACCTTAAGAGTACA
CAAGAAGCTATAAATAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAACCTTCAAAGACTAAG
TGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAA
GCTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGACGAGCATCTATTGGCACTT
GAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGACATAGGAAACGGATGCTTCGAAACCAAACACAAATG
CAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCAT
TGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCT
TCTAGTTTGGCTGTAACATTAATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCCAT
CTGTCTATAA

Fig.36B: SEQ ID NO : 102

Trình tự axit amin của HA B Wisconsin (PrL-).

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTRTRGKLCPDCLNCTD
LDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDAEKAPGGPYRLGT
SGSCPNATSKIGFFATMAWAVPKDNYKNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMKSLYGDSNPQKFTSSA
NGVTTHYVSQIGDFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEAD
CLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKST
QEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLAL
ERKLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNHTILLYYSTAA
SSLAVTLMLAIFIVYMVSRDNVSCSICL*

Fig. 36C: Biểu diễn theo sơ đồ của cấu trúc số 1445 (2X35S/CPMV HT)

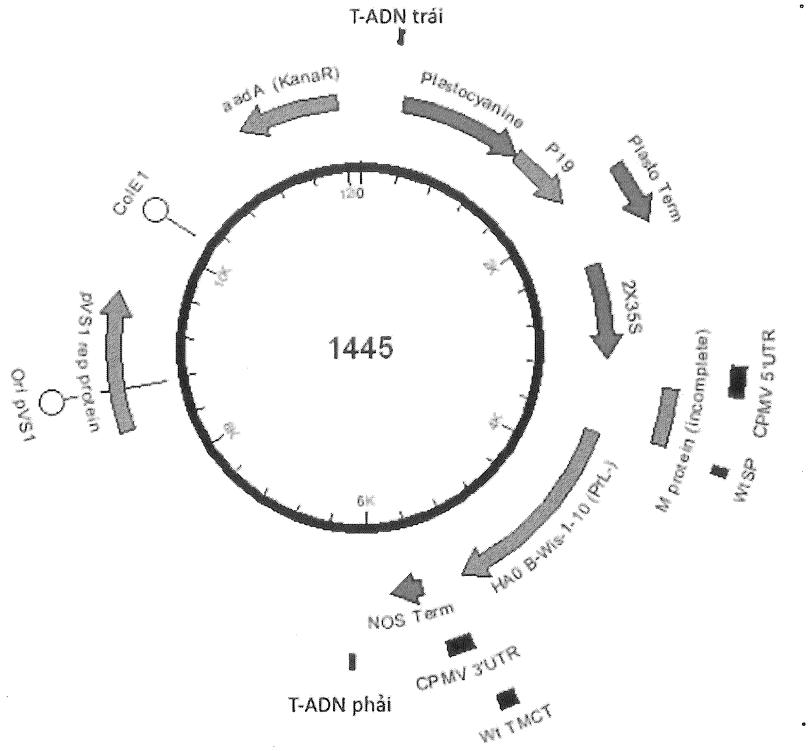


Fig. 36D: Biểu diễn theo sơ đồ của cấu trúc số 1820 (2X35S/CPMV HT* (-Mprot))

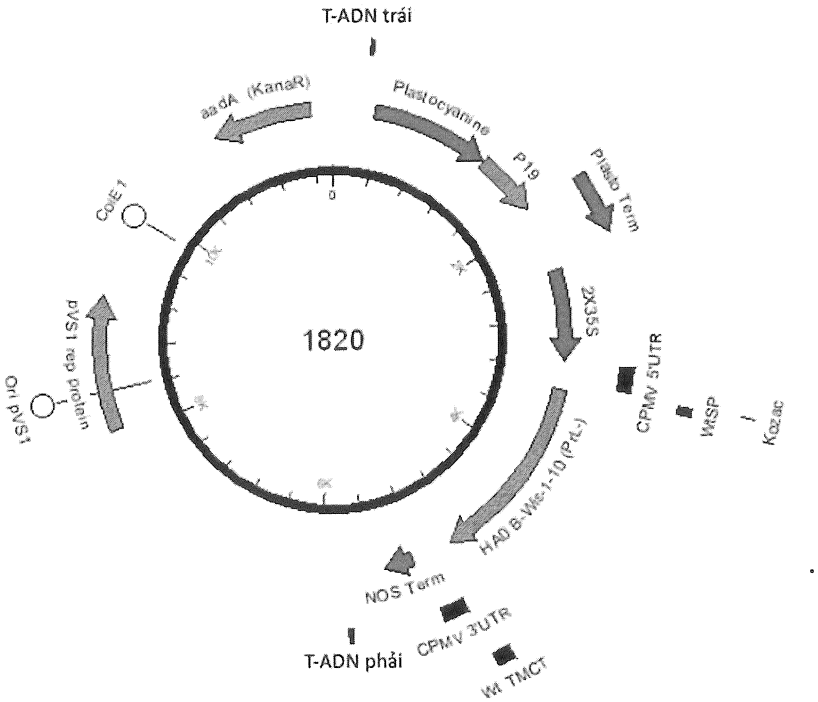
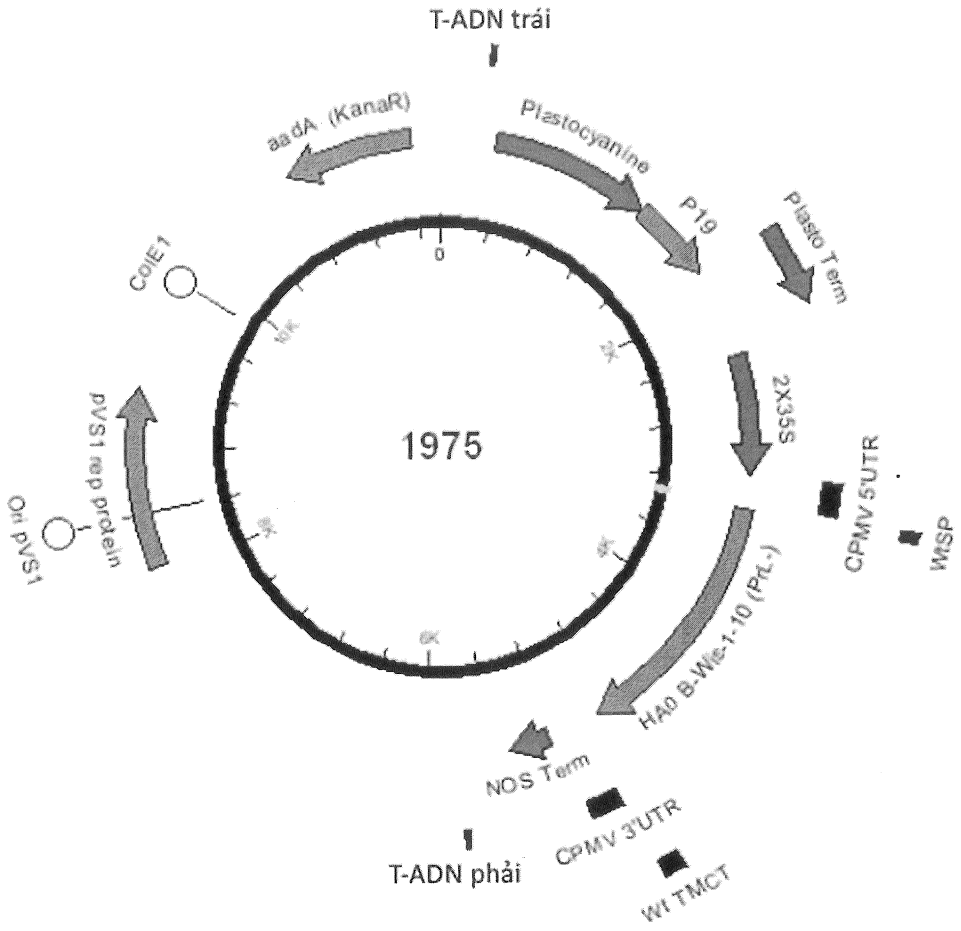


Fig. 36E
Biểu diễn theo sơ đồ của cấu trúc số 1975 (2X35S/CPMV HT* (fl5'UTR))



2X35S/CPMV HT (cấu trúc số 1454) và HT*(-Mprot) (cấu trúc số 1893) đối với HA B Wisconsin (PrL-)+H1 California TMCT

Fig.37A: SEQ ID NO : 103

Trình tự nucleotit của HA B Wisconsin (PrL-)+H1 California TMCT

ATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGGATAACATCTTCAAA
CTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGCGTGATACCACTGACAACAACACCAA
CAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGCCCGGACTGTCTCAACTGTACAGAT
CTGGATGTGGCCTTGGGCAGGCCAATGTGTGTGGGGACCACACCTTCTGCTAAAGCTTCAATACTCCACGAGGTCAG
ACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAAAATCAGGCAACTACCCAATCTTCTCAGAGGAT
ATGAAAATATCAGGTTATCAACCCAAAACGTTATCGATGCAGAAAAAGCACCAGGAGGACCCTACAGACTTGGAACC
TCAGGATCTTGCCCTAACGCTACCAGTAAAATCGGATTTTTTGCAACAATGGCTTGGGCTGTCCCAAAGGACAACTA
CAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAAATTACTGTTTGGG
GGTTCCATTCAGATAACAAAACCCAAATGAAGAGCCTCTATGGAGACTCAAATCCTCAAAAGTTCACCTCATCTGCT
AATGGAGTAACCACACATTATGTTTCTCAGATTGGCGACTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAG
CGGCAGAATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTATCAAAGAGGTGTTTTGT
TGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCATTGCCTTTAATTGGTGAAGCAGAT
TGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAAA
TTGCCCAATATGGGTAAAAACACCTTTGAAGCTTGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAG
GAATGATTGCAGGTTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCAGACCTTAAGAGTACA
CAAGAAGCTATAAATAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAACCTTCAAAGACTAAG
TGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAA
GCTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGACGAGCATCTATTGGCACTT
GAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGACATAGGAAACGGATGCTTCGAAACCAAACACAAATG
CAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCAT
TGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAACTACCAGATTTTGGCGATCTATTCAACTGTCGCC
AGTTCATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAAT
ATGTATTTAA

Fig.37B: SEQ ID NO : 104

Trình tự axit amin của HA B Wisconsin (PrL-)+H1 California TMC.

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTRTRGKLCPDCLNCTD
LDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDAEKAPGGPYRLGT
SGSCPNATSKIGFFATMAWAVPKDNYKNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMKSLYGDSNPQKFTSSA
NGVTTHYVSQIGDFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEAD
CLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKST
QEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLAL
ERKLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNYQILAIYSTVA
SSLVLVVSLGAISFWMCSNGSLQCRICI*

Fig. 37C
Biểu diễn theo sơ đồ của cấu trúc số 1454 (2X35S/CPMV HT)

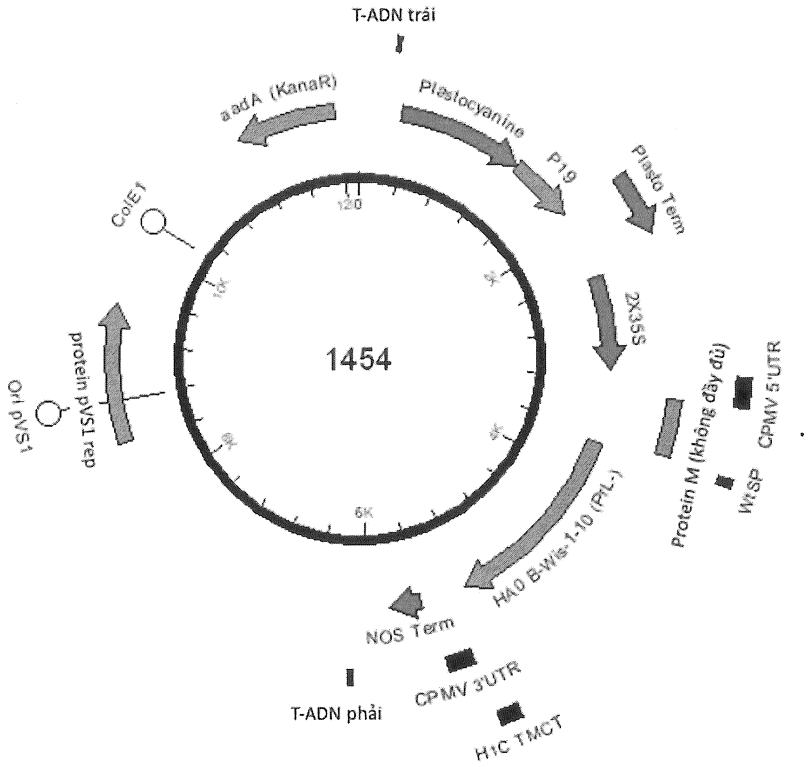
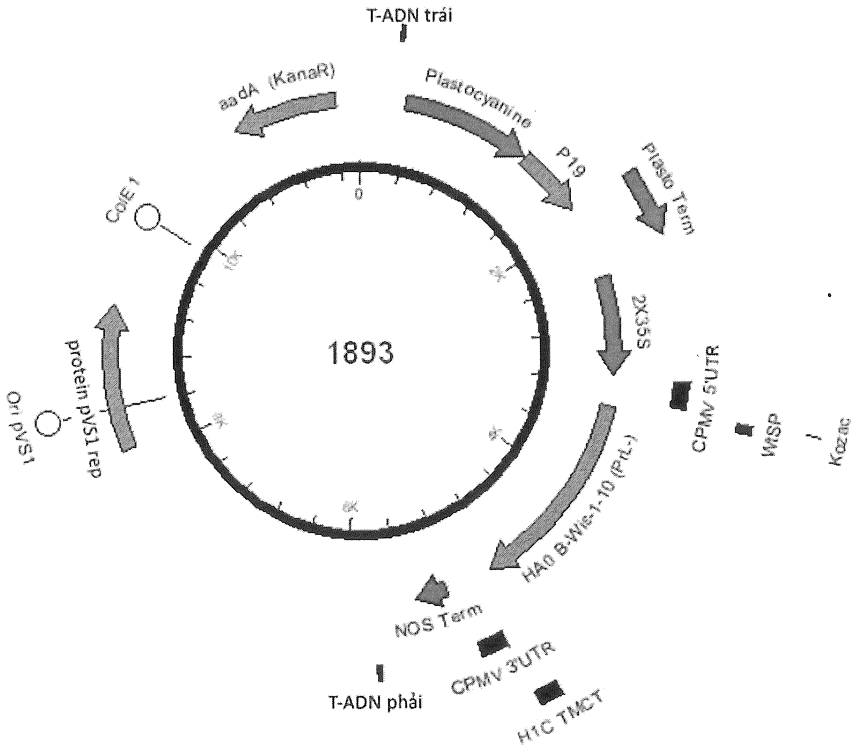


Fig. 37D
Biểu diễn theo sơ đồ của cấu trúc số 1893 (2X35S/CPMV HT* (-Mprot))



L-2X35S/CPMV HT (cấu trúc số 1067) và HT* (cấu trúc số 1875) đối với PDISP/HA B Brishane (PrL-)+H1 Califonia TMCT

Fig.38A: SEQ ID NO : 105

Trình tự nucleotit của PDISP/HA B Brisbane (PrL-)+H1 California TMCT.

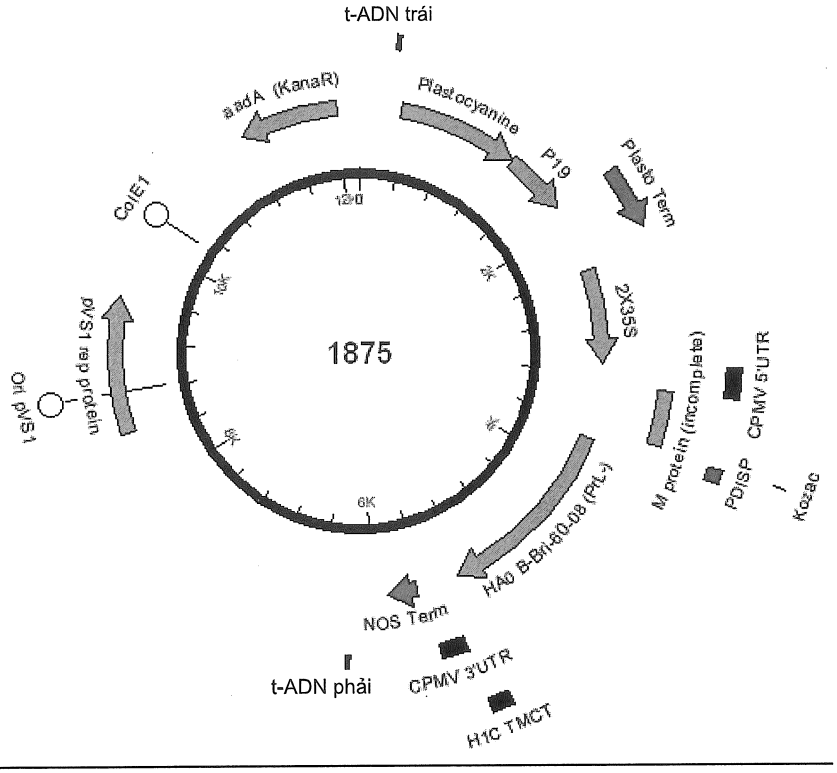
ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCGA
ATCTGCACTGGAATAACATCGTCAAACTCACCACATGTCGTCAAAACTGCTACTCAAGGGGAGGTCAATGTGACTG
GTGTAATACCACTGACAACAACACCCACCAAATCTCATTTTGCAAATCTCAAAGGAACAGAAACCAGGGGGAAAC
TATGCCCAAAATGCCTCAACTGCACAGATCTGGACGTAGCCTTGGGCAGACCAAAATGCACGGGGAAAATACCCT
CGGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTACATCTGGGTGCTTTCCTATAATGCACGACAGAACAAA
AATTAGACAGCTGCCTAACCTTCTCCGAGGATACGAACATATCAGGTTATCAACCCATAACGTTATCAATGCAGAA
AATGCACCAGGAGGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACATTACCAATGGAAACGGATTTTTCG
CAACAATGGCTTGGGCCGTCCCAAAAAACGACAAAAACAAAACAGCAACAAATCCATTAACAATAGAAGTACCA
TACATTTGTACAGAAGGAGAAGACCAAATTACCGTTTGGGGGTTCCACTCTGACAACGAGACCCAAATGGCAAAG
CTCTATGGGGACTCAAAGCCCCAGAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGTTTCACAGATTG
GTGGCTTCCCAAATCAAACAGAAGACGGAGGACTACCACAAAGTGGTAGAATTGTTGTTGATTACATGGTGCAAA
AATCTGGGAAAACAGGAACAATTACCTATCAAAGGGGTATTTTATTGCCTCAAAAGGTGTGGTGCGCAAGTGGCA
GGAGCAAGGTAATAAAAGGATCCTTGCCTTTAATTGGAGAAGCAGATTGCCTCCACGAAAAATACGGTGGATTAA
ACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAGGAAATTGCCCAATATGGGTGAAAACACCCT
TGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGAT
ACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAACAAGATA
ACAAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCCATGGATGAACTC
CACAACGAAATACTAGAACTAGATGAGAAAGTGGATGATCTCAGAGCTGATACAATAAGCTCACAAATAGAACTC
GCAGTCCTGCTTTCCAATGAAGGAATAATAAACAGTGAAGATGAACATCTCTTGGCGCTTGAAAGAAAGCTGAAG
AAAATGCTGGGCCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTGAAACCAAACACAAGTGCAACCAGACCTGT
CTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTCTCCCCACCTTTGATTCACTGAATATTACTGC
TGCATCTTTAAATGACGATGGATTGGATAATTACCAGATTTTGGCGATCTATTCAACTGTCGCCAGTTCATTGGTAC
TGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAA

Fig.38B: SEQ ID NO : 106

Trình tự axit amin của PDISP/HA B Brisbane (PrL-)+H1 California TMCT.

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKGTETRGKLCPKC
LNCTDLDVALGRPKCTGKIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYEHIRLSTHNVINAENAPGGPYKIG
TSGSCPNITNGNGFFATMAWAVPKNDKNKTATNPLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSA
NGVTTHYVSQIGGFPNQTEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCL
HEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQ
EAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKK
MLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSLNITAASLNDDGLDNYQILAIYSTVASSLVLVVSL
GAISFWMCSNGSLQCRICI*

Fig.38C
Sự biểu diễn theo sơ đồ của cấu trúc số 1875 (2X35S/CPMV HT*)



M-2X35S/CPMV HT (cấu trúc số 2072) và HT* (cấu trúc số 2052) đối với PDISP/HA B Massachussetts (PrL-)

Fig.39A: SEQ ID NO : 107

Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-).

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCGA
ATCTGCACTGGGATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTG
GTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAAGACCAGAGGGAAAC
TATGCCCAGACTGTCTCAACTGTACAGATCTGGATGTGGCCCTGGGCAGGCCAATGTGTGTGGGAACTACACCTTC
TGCGAAAGCTTCAATACTTCACGAAGTCAGACCTGTTACATCCGGGTGCTTCCCTATAATGCACGACAGAACAAAA
ATCAGGCAACTAGCCAATCTTCTCAGAGGATATGAAAATATCAGGTTATCAACCCAAAACGTTATCGATGCAGAA
AAGGCACCAGGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAAAGCGGATTTTTC
GCAACAATGGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCATTAACAGTAGAAGTACCATAC
ATTTGTGCAGAAGGGGAAGACCAAATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTC
TATGGAGACTCAAATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACATTATGTTTCTCAGATTGGCG
GCTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAGCGGCAGAATTGTCGTTGATTACATGATGCAAAAAC
CTGGGAAAACAGGAACAATTGTCTATCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGA
GCAAAGTAATAAAAGGGTCCTTGCCTTTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACA
AAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTGAAAACACCTTTGA
AGCTTGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACA
CATCTCACGGAGCACATGGAGTGGCAGTTGCTGCAGACCTTAAGAGCACACAAGAAGCTATAAACAAGATAACAA
AAAATCTCAACTCTTTGAGTGAGCTAGAAGTAAAGAATCTTCAAAGGCTAAGTGGTGCCATGGATGAACTCCACA
ACGAAATACTCGAGCTGGATGAGAAAGTGGATGACCTCAGAGCTGACACTATAAGTTCACAAATAGAACTTGCAG
TCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAA
TGCTGGGTCCCTCTGCTGTAGACATAGGAAATGGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTAG
ACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAGTTTTCTCTCCCCACTTTTGATTCATTGAACATTACTGCTGCA
TCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTTCTAGTTTGGCTGTAAC
ATTGATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCCATCTGTCTATAA

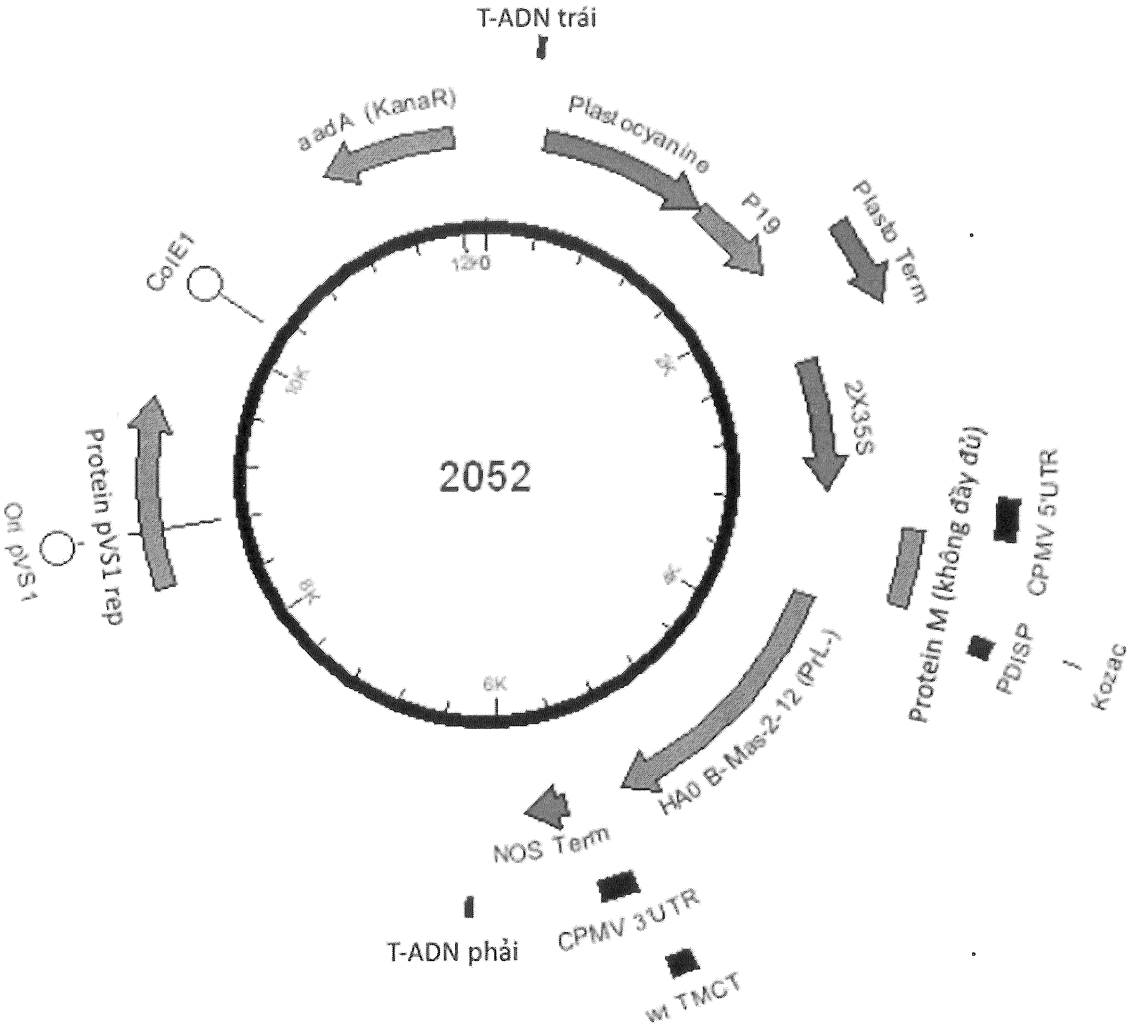
Fig.39B: SEQ ID NO :108

Trình tự axit amin của PDISP/HA B Massachussetts (PrL-).

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTKTRGKLCPDC
LNCTDLDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLANLLRGYENIRLSTQNVIDAEKAPGGPYRL
GTSGSCPNATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICAEGEDQITVWGFHSDNKTQMKNLYGDSNPQKFTSS
ANGVTTHYVSQIGGFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEAD
CLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKS
TQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKL
KKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNHTILLYYSTAASSLAVTL
MLAIFIVYMVSRDNVSCSICL*

Fig. 39C

Biểu diễn theo sơ đồ của cấu trúc số 2052 (2X35S/CPMV HT*)



N-2X35S/CPMV HT (cấu trúc số 2074) và HT* (cấu trúc số 2062) đối với PDISP/HA B Massachussetts (PrL-)+H1 California TMCT

Fig.40A: SEQ ID NO : 109

Trình tự nucleotit của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT.

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCGA
ATCTGCACTGGGATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTG
GTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAAGACCAGAGGGAAAC
TATGCCCAGACTGTCTCAACTGTACAGATCTGGATGTGGCCCTGGGCAGGCCAATGTGTGTGGGAACTACACCTTC
TGCGAAAGCTTCAATACTTCACGAAGTCAGACCTGTTACATCCGGGTGCTTCCCTATAATGCACGACAGAACAAAA
ATCAGGCAACTAGCCAATCTTCTCAGAGGATATGAAAATATCAGGTTATCAACCCAAAACGTTATCGATGCAGAA
AAGGCACCAGGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAAAGCGGATTTTTC
GCAACAATGGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCATTAACAGTAGAAGTACCATAC
ATTTGTGCAGAAGGGGAAGACCAAATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTC
TATGGAGACTCAAATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACATTATGTTTCTCAGATTGGCG
GCTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAGCGGCAGAATTGTCGTTGATTACATGATGCAAAAAC
CTGGGAAAACAGGAACAATTGTCTATCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGA
GCAAAGTAATAAAAGGGTCCTTGCCTTTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACA
AAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTGAAAACACCTTTGA
AGCTTGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACA
CATCTCACGGAGCACATGGAGTGGCAGTTGCTGCAGACCTTAAGAGCACACAAGAAGCTATAAACAAGATAACAA
AAAATCTCAACTCTTTGAGTGAGCTAGAAGTAAAGAATCTTCAAAGGCTAAGTGGTGCCATGGATGAACTCCACA
ACGAAATACTCGAGCTGGATGAGAAAGTGGATGACCTCAGAGCTGACACTATAAGTTCACAAATAGAACTTGCAG
TCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAA
TGCTGGGTCCCTCTGCTGTAGACATAGGAAATGGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTAG
ACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAGTTTTCTCTCCCCACTTTTGATTCATTGAACATTACTGCTGCA
TCTTTAAATGATGATGGATTGGATAACTACCAGATTTTGGCGATCTATTCAACTGTCGCCAGTTCATTGGTACTGGT
AGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAA

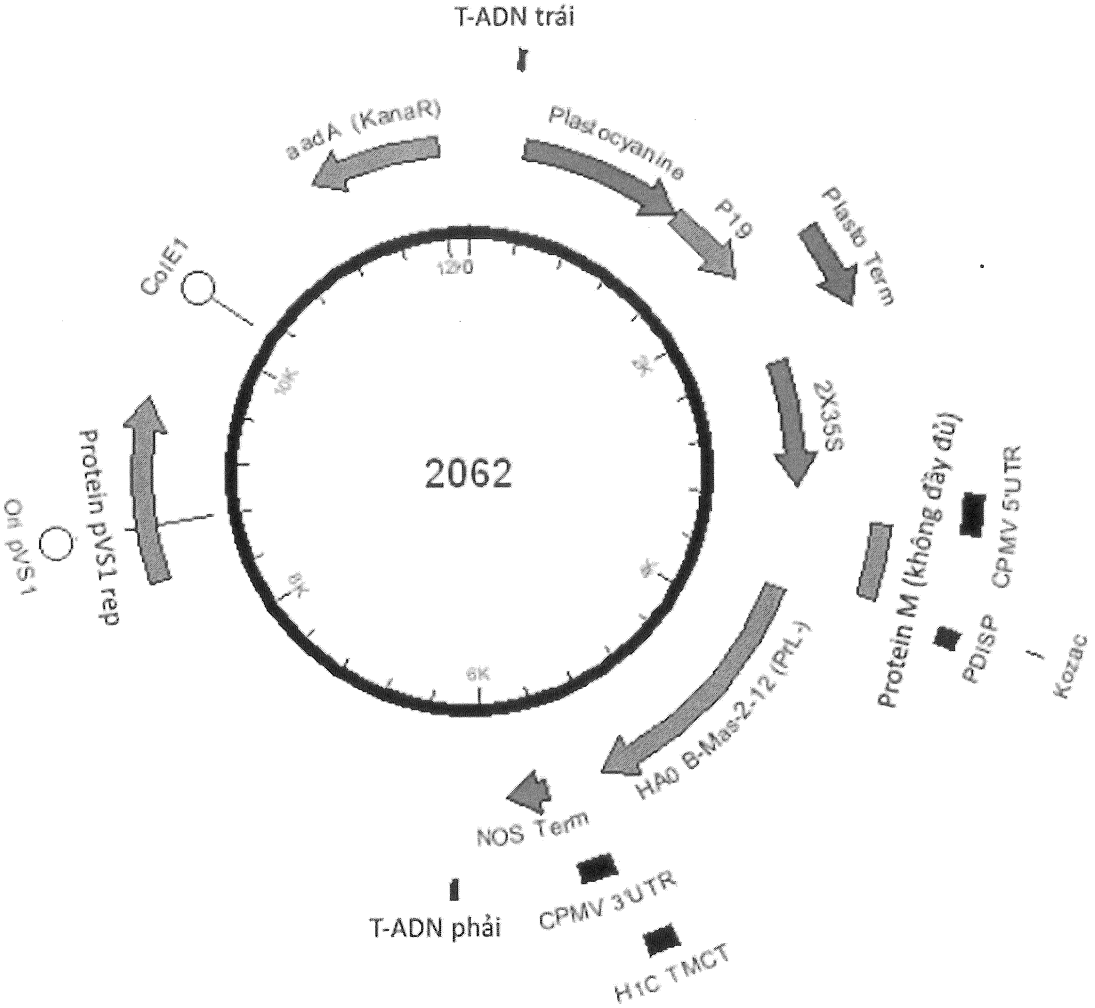
Fig.40B: SEQ ID NO : 110

Trình tự axit amin của PDISP/HA B Massachussetts (PrL-)+H1 California TMCT.

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTKTRGKLCPDC
LNCTDLDVALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLANLLRGYENIRLSTQNVIDAEKAPGGPYRL
GTSGSCPNATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICAEGEDQITVWGFHSDNKTQMKNLYGDSNPQKFTSS
ANGVTTHYVSQIGGFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEAD
CLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKS
TQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKL
KKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNYQILAIYSTVASSLVLVV
SLGAISFWMCSNGSLQCRICI*

Fig. 40C

Biểu diễn theo sơ đồ của cấu trúc số 2062 (2X35S/CPMV HT*)



O-2X35S/CPMV HT (cấu trúc số 1445) và HT* (cấu trúc số 1839), HT*(-Mprot) (cấu trúc số 1820) và HT(fl5'UTR) (cấu trúc số 1975) đối với HA B Wisconsin (PrL-)

Fig.41A: SEQ ID NO :111

Trình tự nucleotit của HA B Wisconsin (PrL-).

ATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGGATAACATCTTCAA
ACTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGCGTGATACCACTGACAACAACAC
CAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGCCCGGACTGTCTCAACTGTA
CAGATCTGGATGTGGCCTTGGGCAGGCCAATGTGTGTGGGGACCACACCTTCTGCTAAAGCTTCAATACTCCACGA
GGTCAGACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAAAATCAGGCAACTACCCAATCTTCTC
AGAGGATATGAAAATATCAGGTTATCAACCCAAAACGTTATCGATGCAGAAAAAGCACCAGGAGGACCCTACAGA
CTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAAATCGGATTTTTTGCAACAATGGCTTGGGCTGTCCCAA
AGGACAACTACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAA
ATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAGCCTCTATGGAGACTCAAATCCTCAAAAGT
TCACCTCATCTGCTAATGGAGTAACCACACATTATGTTTCTCAGATTGGCGACTTCCCAGATCAAACAGAAGACGG
AGGACTACCACAAAGCGGCAGAATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTA
TCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCATTGCC
TTTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGA
ACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTAAAAACACCTTTGAAGCTTGCCAATGGAACCAAATATAG
ACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCTCACGGAGCACATGGAGTGGC
AGTGGCGGCAGACCTTAAGAGTACACAAGAAGCTATAAATAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCT
AGAAGTAAAGAACCTTCAAAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAA
AGTGGATGATCTCAGAGCTGACACTATAAGCTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATA
AACAGTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGACATA
GGAAACGGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAAT
GCAGGAGAATTTTCTCTCCCCACTTTTGATTCATTGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAA
CCATACTATACTGCTCTATTACTCAACTGCTGCTTCTAGTTTGGCTGTAACATTAATGCTAGCTATTTTTATTGTTTA
TATGGTCTCCAGAGACAACGTTTCATGCTCCATCTGTCTATAA

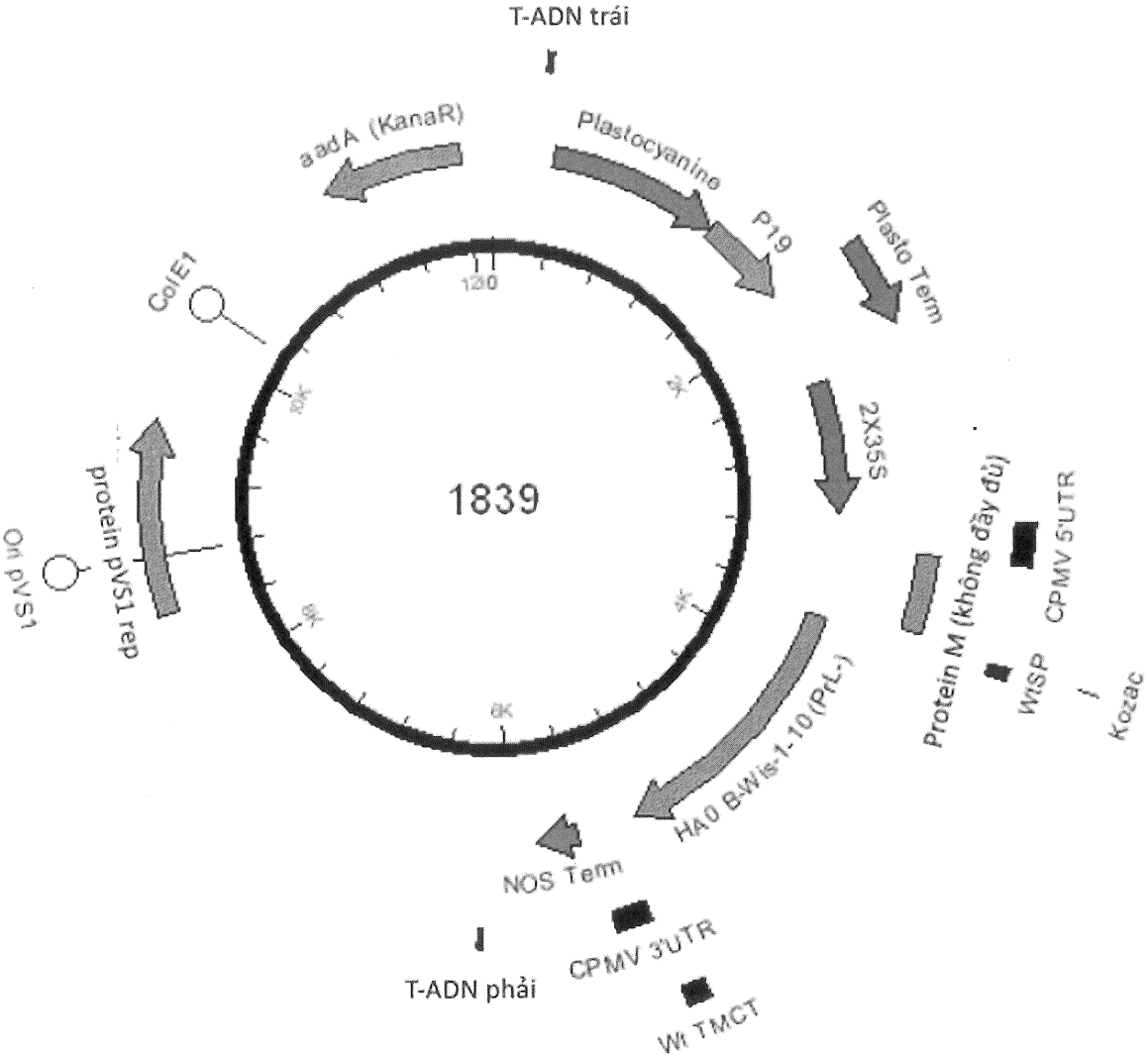
Fig.41B: SEQ ID NO : 112

Trình tự axit amin của HA B Wisconsin (PrL-).

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTRTRGKLCPDCLNCTDLDV
ALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDAEKAPGGPYRLGTSGSCPN
ATSKIGFFATMAWAVPKDNYKNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMKSLYGDSNPQKFTSSANGVTTHY
VSQIGDFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGG
LNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKIT
KNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSA
VDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNHTILLYYSTAASSLAVTLMLAIFIVY
MVSRDNVSCSICL*

Fig. 41C

Biểu diễn theo sơ đồ của cấu trúc số 1839 (2X35S/CPMV HT*)



P-2X35S/CPMV HT (cấu trúc số 1454) và HT* (cấu trúc số 1860) đối với HA B Wisconsin (PrL-)+H1 California TMCT

Fig.42A: SEQ ID NO : 113

Trình tự nucleotit của HA B Wisconsin (PrL-)+H1 California TMCT

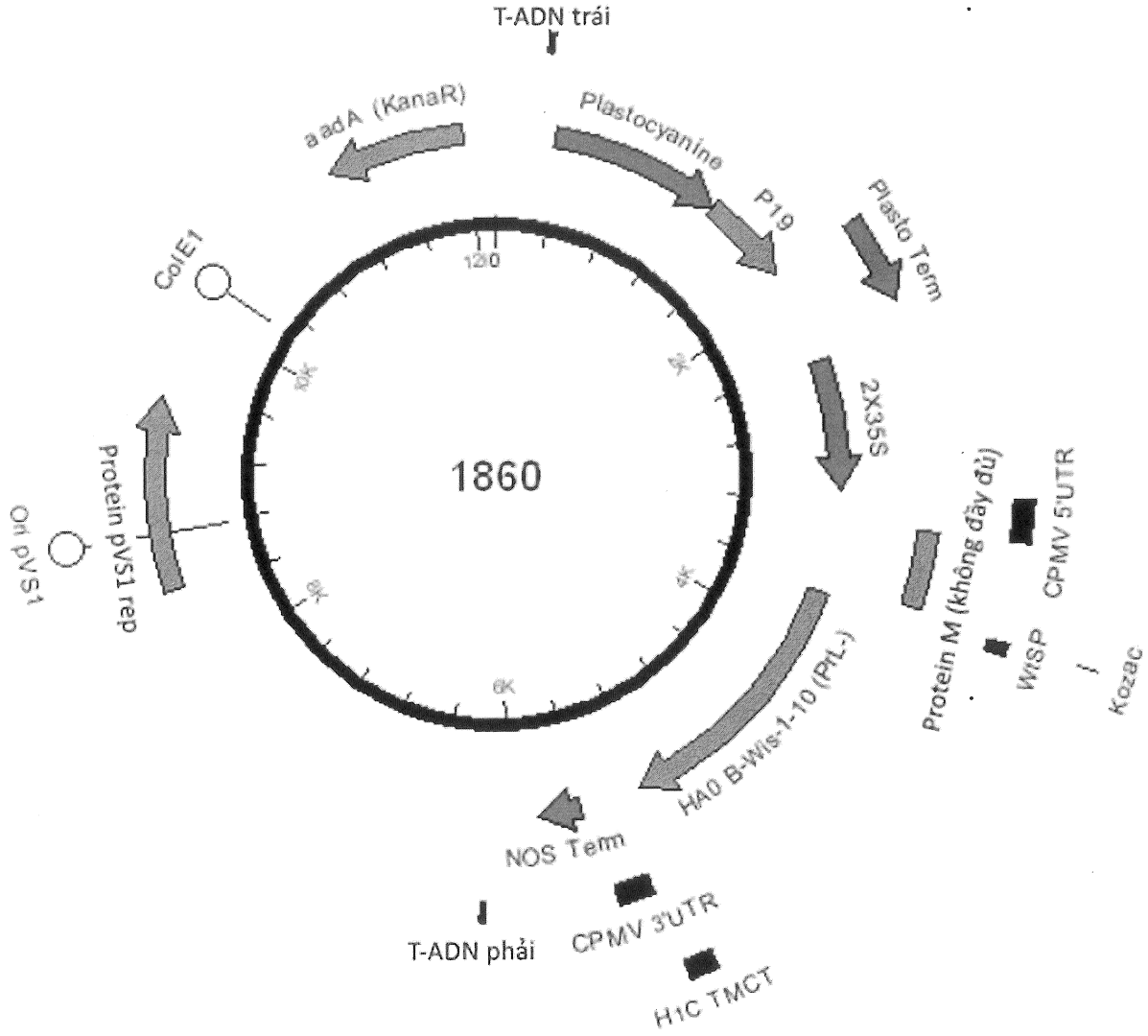
ATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAATCTGCACTGGGATAACATCTTCAA
ACTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGCGTGATACCACTGACAACAACAC
CAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGCCCGGACTGTCTCAACTGTA
CAGATCTGGATGTGGCCTTGGGCAGGCCAATGTGTGTGGGGACCACACCTTCTGCTAAAGCTTCAATACTCCACGA
GGTCAGACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAAAATCAGGCAACTACCCAATCTTCTC
AGAGGATATGAAAATATCAGGTTATCAACCCAAAACGTTATCGATGCAGAAAAAGCACCAGGAGGACCCTACAGA
CTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAAATCGGATTTTTTGCAACAATGGCTTGGGCTGTCCCAA
AGGACAACTACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAA
ATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAGCCTCTATGGAGACTCAAATCCTCAAAAGT
TCACCTCATCTGCTAATGGAGTAACCACACATTATGTTTCTCAGATTGGCGACTTCCCAGATCAAACAGAAGACGG
AGGACTACCACAAAGCGGCAGAATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTA
TCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCATTGCC
TTTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGA
ACATGCAAAAGCCATAGGAAATTGCCCAATATGGGTAAAAACACCTTTGAAGCTTGCCAATGGAACCAAATATAG
ACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCTCACGGAGCACATGGAGTGGC
AGTGGCGGCAGACCTTAAGAGTACACAAGAAGCTATAAATAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCT
AGAAGTAAAGAACCTTCAAAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAA
AGTGGATGATCTCAGAGCTGACACTATAAGCTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATA
AACAGTGAAGACGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGACATA
GGAAACGGATGCTTCGAAACCAAACACAAATGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAAT
GCAGGAGAATTTTCTCTCCCCACTTTTGATTCATTGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAA
CTACCAGATTTTGGCGATCTATTCAACTGTCGCCAGTTCATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCT
GGATGTGCTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAA

Fig.42B: SEQ ID NO : 114

Trình tự axit amin của HA B Wisconsin (PrL-)+H1 California TMC.

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTRTRGKLCPDCLNCTDLDV
ALGRPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDAEKAPGGPYRLGTSGSCPN
ATSKIGFFATMAWAVPKDNYKNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMKSLYGDSNPQKFTSSANGVTTHY
VSQIGDFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGG
LNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKIT
KNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSA
VDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNYQILAIYSTVASSLVLVVSLGAISFW
MCSNGSLQCRICI*

Fig. 42C
Biểu diễn theo sơ đồ của cấu trúc số 1860 (2X35S/CPMV HT*)



2X35S/CPMV HT (cấu trúc số 489), HT*(-Mprot) (cấu trúc số 1880) v HT(fl5'UTR) (cấu trúc số 1885) đối với H5 Indonesia

Fig.43A:

Trình tự nucleotit của H5 Indonesia nguyên thể. (SEQ ID NO: 115)

ATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCAGTCTTGTTAAAAGTGATCAGATTTGCATTGGTTACCATGCAAA
CAATTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGTTACTGTTACACATGCCCAAGACATACTGGAAAAGA
CACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTTAAGAGATTGTAGTGTAGCTGGATGGCTC
CTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCGGAATGGTCTTACATAGTGGAGAAGGCCAATCCAACCAA
TGACCTCTGTTACCCAGGGAGTTTCAACGACTATGAAGAACTGAAACACCTATTGAGCAGAATAAACCATTTTGAGA
AAATTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCTCATCAGGAGTTAGCTCAGCATGTCCATACCTG
GGAAGTCCCTCCTTTTTTAGAAATGTGGTATGGCTTATCAAAAAGAACAGTACATACCCAACAATAAAGAAAAGCTA
CAATAATACCAACCAAGAGGATCTTTTGGTACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGC
TATATCAAAACCCAACCACCTATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAAAATAGCTACT
AGATCCAAAGTAAACGGGCAAAGTGGAAGGATGGAGTTCTTCTGGACAATTTTAAAACCTAATGATGCAATCAACTT
CGAGAGTAATGGAAATTTCATTGCTCCAGAATATGCATACAAAATTGTCAAGAAAGGGGACTCAGCAATTATGAAAA
GTGAATTGGAATATGGTAACTGCAACACCAAGTGTCAAACTCCAATGGGGGCGATAAACTCTAGTATGCCATTCCAC
AACATACACCCTCTCACCATCGGGGAATGCCCCAAATATGTGAAATCAAACAGATTAGTCCTTGCAACAGGGCTCAG
AAATAGCCCTCAAAGAGAGAGCAGAAGAAAAAAGAGAGGACTATTTGGAGCTATAGCAGGTTTTATAGAGGGAGGAT
GGCAGGGAATGGTAGATGGTTGGTATGGGTACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAA
TCCACTCAAAAGGCAATAGATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGC
CGTTGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACGGGTTTCTAGATG
TCTGGACTTATAATGCCGAACTTCTGGTTCTCATGGAAAATGAGAGAACTCTAGACTTTCATGACTCAAATGTTAAG
AACCTCTACGACAAGGTCCGACTACAGCTTAGGGATAATGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATCA
CAAATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAACTATCCGCAGTATTCAGAAGAAGCAAGAT
TAAAAAGAGAGGAAATAAGTGGGGTAAAATTGGAATCAATAGGAACTTACCAAATACTGTCAATTTATTCAACAGTG
GCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAG
AATTTGCATTTAA

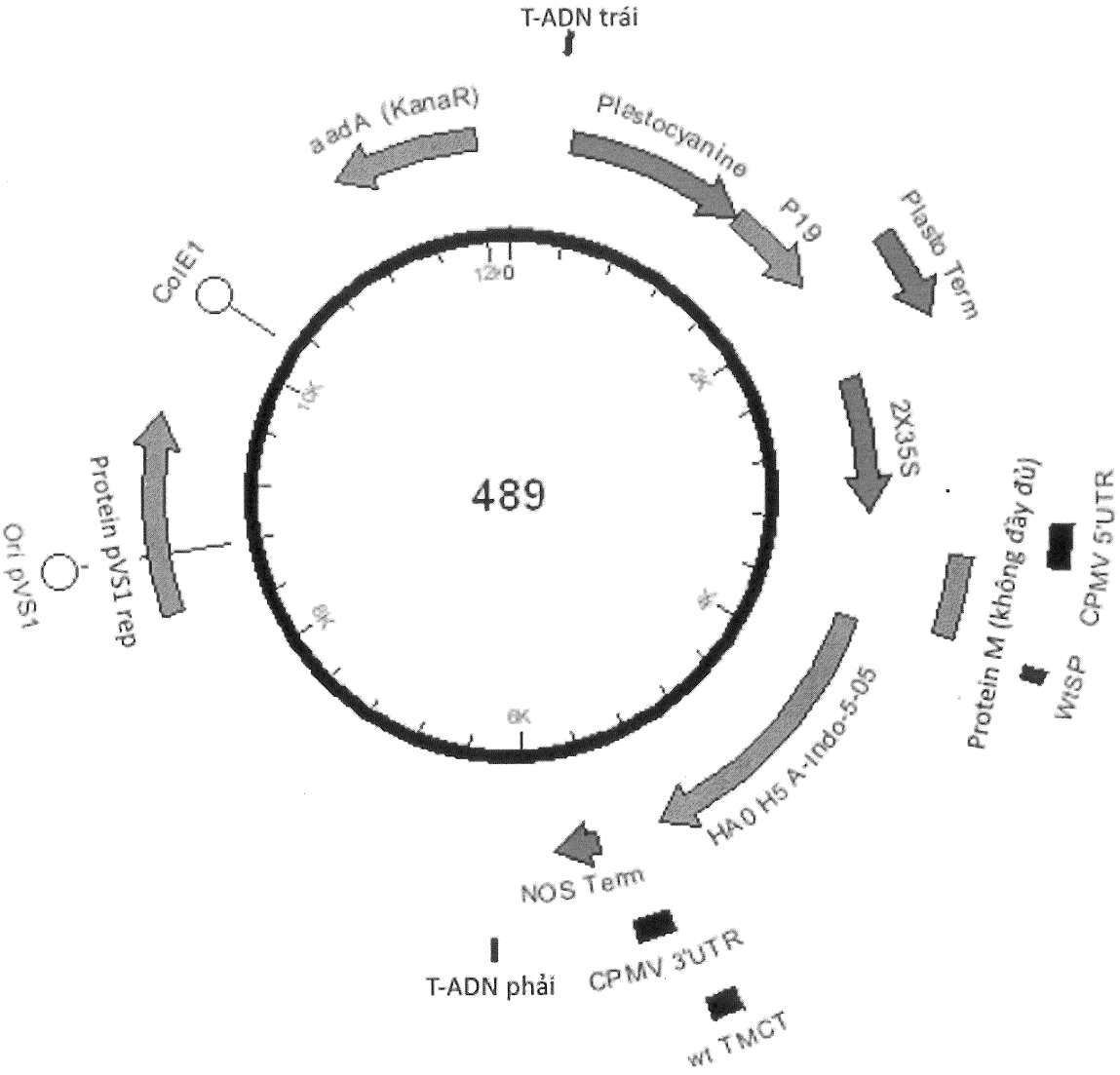
Fig.43B:

Trình tự axit amin của H5 Indonesia nguyên thể. (SEQ ID NO: 116)

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCDLDGVKPLILRDCSVAGWL
LGNPMCDEFINVPEWSYIVEKANPTNDLCYPGSFNDYEELKHLLSRINHFEKIQIIPKSSWSDHEASSGVSSACPYL
GSPSFFRNVVWLIKKNSTYPTIKKSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISIGTSTLNQRLVPKIAT
RSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSAIMKSELEYGNCNTKCQTPMGAINSSMPFH
NIHPLTIGECPKYVKSNRLVLATGLRNSPQRESRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKE
STQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVK
NLYDKVRLQLRDNAKELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEISGVKLESIGTYQILSIYSTV
ASSLALAIMMAGLSLWMCSNGSLQCRICI*

Fig. 43C

Biểu diễn theo sơ đồ của cấu trúc số 489 (2X35S/CPMV HT)



2X35S/CPMV-HT*(-Mprot)/ PDISP/H3 Victoria/ NOS (Cấu trúc số 1800)

Fig.44A:
IF**(SacII)-PDI.s1+4c
ACAGGGCCCAATACCGCGGAGAAAATGGCGAAAAACGTTGCGATTTTCGGCT (SEQ ID NO: 117)

Fig.44B:
IF-H3V36111.s1-4r
ACTAAAGAAAATAGGCCTTCAAATGCAAATGTTGCACCTAATGTTGCCCTT (SEQ ID NO: 118)

Fig.44C:
Trình tự nucleotit của PDISP/H3 Victoria. (SEQ ID NO: 119)
ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCCAAAA
ACTTCCTGGAAATGACAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATAGTGAAAA
CAATCACGAATGACCAAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAATTCCTCAATAGGTGAAATATGCGAC
AGTCCTCATCAGATCCTTGATGGAGAAAACTGCACACTAATAGATGCTCTATTGGGAGACCCTCAGTGTGATGGCTT
CCAAAATAAGAAATGGGACCTTTTTGTTGAACGAAGCAAAGCCTACAGCAACTGTTACCCTTATGATGTGCCGGATT
ATGCCTCCCTTAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTTTAACAATGAAAGCTTCAATTGGACTGGAGTC
ACTCAAAACGGAACAAGTTCTGCTTGCATAAGGAGATCTAATAATAGTTTCTTTAGTAGATTAAATTGGTTGACCCA
CTTAAACTTCAAATACCCAGCATTGAACGTGACTATGCCAAACAATGAACAATTTGACAAATTGTACATTTGGGGGG
TTCACCACCCGGGTACGGACAAGGACCAAATCTTCCTGTATGCTCAATCATCAGGAAGAATCACAGTATCTACCAAA
AGAAGCCAACAAGCTGTAATCCCGAATATCGGATCTAGACCCAGAATAAGGAATATCCCTAGCAGAATAAGCATCTA
TTGGACAATAGTAAAACCGGGAGACATACTTTTGATTAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCA
AAATACGAAGTGGGAAAAGCTCAATAATGAGATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCA
AATGGAAGCATTCCCAATGACAAACCATTCCAAAATGTAAACAGGATCACATACGGGGCCTGTCCCAGATATGTTAA
GCAAAGCACTCTGAAATTGGCAACAGGAATGCGAAATGTACCAGAGAAACAAACTAGAGGCATATTTGGCGCAATAG
CGGGTTTCATAGAAAATGGTTGGGAGGGAATGGTGGATGGTTGGTACGGTTTCAGGCATCAAAATTCTGAGGGAAGA
GGACAAGCAGCAGATCTCAAAAGCACTCAAGCAGCAATCGATCAAATCAATGGGAAGCTGAATCGATTGATCGGGAA
AACCAACGAGAAATTCCATCAGATTGAAAAAGAATTCTCAGAAGTCGAAGGGAGAATTCAGGACCTTGAGAAATATG
TTGAGGACACTAAAATAGATCTCTGGTCATACAACGCGGAGCTTCTTGTTGCCCTGGAGAACCAACATACAATTGAT
CTAACTGACTCAGAAATGAACAAACTGTTTGAAAAAACAAAGAAGCAACTAAGGGAAAATGCTGAGGATATGGGCAA
TGGTTGTTTCAAAATATACCACAAATGTGACAATGCCTGCATAGGATCAATCAGAAATGGAACTTATGACCACGATG
TATACAGAGATGAAGCATTAAACAACCGGTTCCAGATCAAGGGAGTTGAGCTGAAGTCAGGGTACAAAGATTGGATC
CTATGGATTTCCTTTGCCATATCATGTTTTTTGCTTTGTGTTGCTTTGTTGGGGTTCATCATGTGGGCCTGCCAAAA
GGGCAACATTAGGTGCAACATTTGCATTTGA

Fig. 44D

Biểu diễn theo sơ đồ của cấu trúc 2171. Vị trí enzym giới hạn SacII và StuI được sử dụng để duỗi thẳng plasmit được chú giải trên sơ đồ biểu diễn

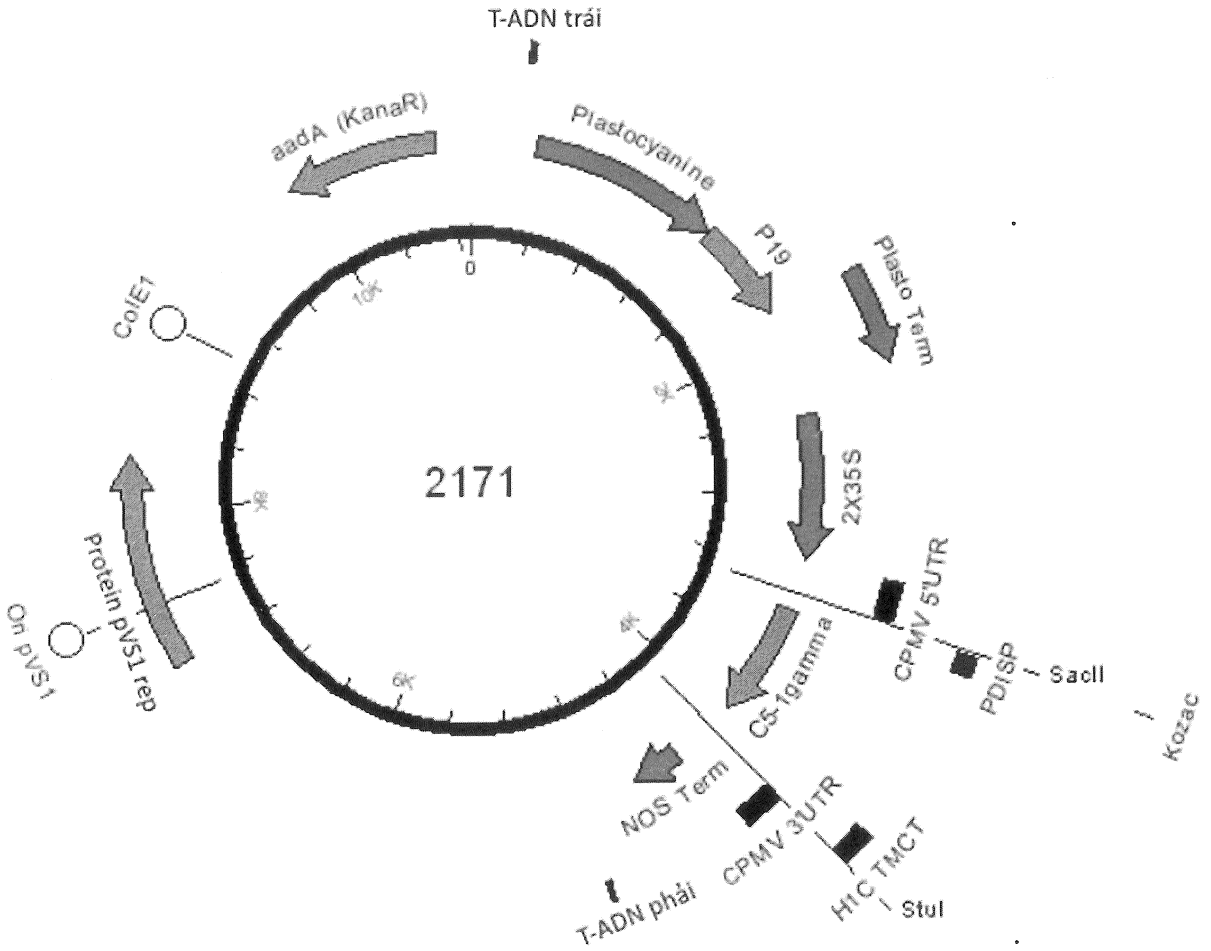


Fig.44E
Cấu trúc 2171 từ bên trái sang bên phải đường biên t-ADN (gạch dưới).
2X35S/CPMV-HT*(-Mprot)/NOS với Plastoxyanin-P19-Plastoxyanin làm cảm cát xét biểu hiện chất ức chế (SEQ ID NO:120)

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACTTAATAACACATTGCGGACGTTTTTAATGTACT
GAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAAAGTTTAAGTTAGCA
AGTGTGTACATTTTTACTTGAACAAAAATATTCACCTACTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAA
ATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTTGCAACATTTGAGAAAATTTTGTTGTTCT
CTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGAG
AAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGG
TTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTAGAGAATTTTTGGCAAGTCATTAAAAAG
AAAGAATAAATTATTTTTAAAATTAAAAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGA
GTTGGATTAAAGTTGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGC
CCCATAGAGTCAGTTAACTCATTTTTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAA
AAAAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATCCA
ATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACATCTACATTATCTAA
ATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACA
CTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAG
AGAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGAACAAGCTAACAGTGAACGTTGGGATGGAGGATCAGGA
GGTACCACTTCTCCCTTCAAACTTCCTGACGAAAGTCCGAGTTGGACTGAGTGGCGGCTACATAACGATGAGACGAA
TTCGAATCAAGATAATCCCCTTGGTTTCAAGGAAAGCTGGGGTTTCGGGAAAGTTGTATTTAAGAGATATCTCAGAT
ACGACAGGACGGAAGCTTCACTGCACAGAGTCCTTGGATCTTGGACGGGAGATTCGGTTAACTATGCAGCATCTCGA
TTTTTCGGTTTCGACCAGATCGGATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTC
GCGAACTCTTCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGG
AAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAATGCTTCTT
CGTCTCCTATTTATAATATGGTTTGTTATTGTTAATTTTGTTCTTGTAGAAGAGCTTAATTAATCGTTGTTGTTATG
AAATACTATTTGTATGAGATGAACTGGTGTAATGTAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTC
CATAACTAACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATATTTGAACAACTAAAATTGAACATCTTTTGCC
ACAACTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCG
AAATTCATTAACAATCAACTTAACGTTATTAACTACTAATTTTATATCATCCCCTTTGATAAATGATAGTACACCAA
TTAGGAAGGAGCATGCTCGCCTAGGAGATTGTCGTTTCCCGCCTTCAGTTTGCAAGCTGCTCTAGCCGTGTAGCCAA
TACGCAAACCGCCTCTCCCCGCGCGTTGGGAATTACTAGCGCGTGTCGACAAGCTTGCATGCCGGTCAACATGGTGG
AGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAA
CAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAA
GGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTC
CCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGAT
TGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAG
GGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTA
TTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGAT
GCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCAC
GTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAG
ACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAACG
TGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTG
CGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAGGGCCCAATACCGCGGAGAAAATGGCGAAA
AACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGACGTCACTCCTCAG
CCAAAACGACACCCCCATCTGTCTATCCACTGGCCCCTGGATCTGCTGCCCAAACTAACTCCATGGTGACCCTGGGA
TGCCTGGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACTCTGGATCCCTGTCCAGCGGTGTGCACAC
CTTCCCAGCTGTCCTGCAGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGTCCCCTCCAGCACCTGGCCCAGCG
AGACCGTCACCTGCAACGTTGCCCACCCGGCCAGCAGCACCAAGGTGGACAAGAAAATTGTGCCCAGGGATTGTGGT
TGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCATCTTCCCCCCAAAGCCCAAGGATGTGCTCAC
CATTACTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGT
TTGTAGATGATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGCTCAGTC
AGTGAACTTCCCATCATGCACCAGGACTGGCTCAATGGCAAGGAGACGTCCAGATTTTGGCGATCTATTCAACTGTC
GCCAGTTCATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAG
AATATGTATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGG
TTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCA

GCAAGGACACAAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGA
CCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCAT
ATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTA
TGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCG
CGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCCCACGTGACTAGTGGCACTGGCCGT
CGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCA
GCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCTAGAGC
AGCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGACAGGATATATTGGCGGGT
AAACCTAAGAGAAAAGAGCGTTTA

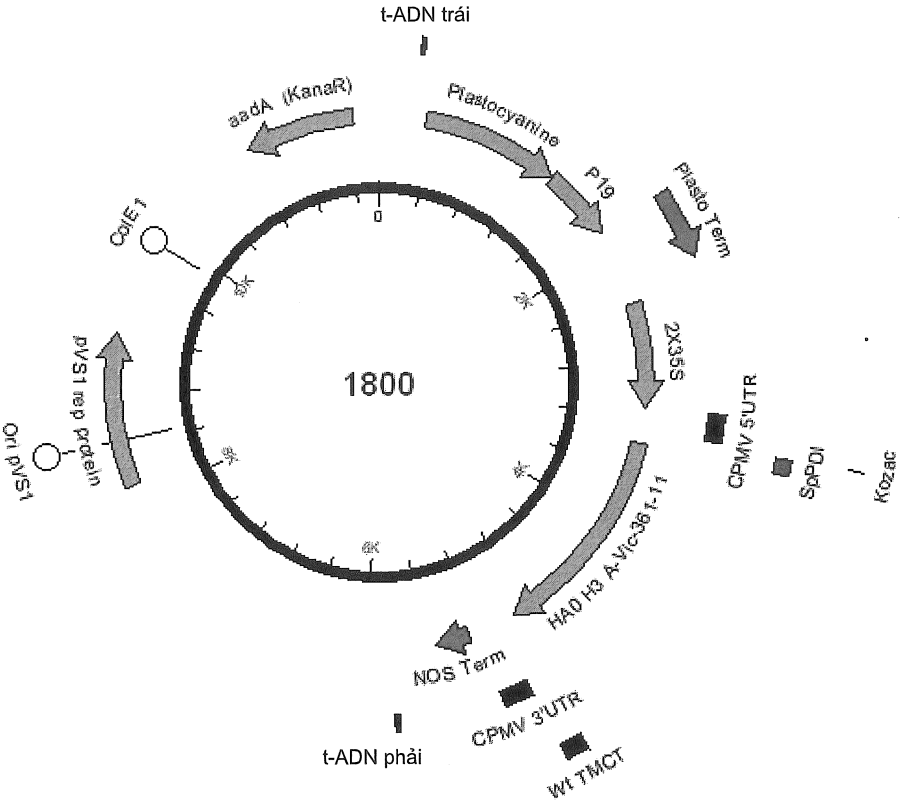
Fig.44F
Cát xét biểu hiện số 1800 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. Trình tự nucleotit PDISP/H3 Victoria được gạch dưới. SEQ ID NO:121

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTT
TCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAG
GTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCC
CCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCA
CGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATAT
CCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGC
CATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCAT
CGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAAT
CCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAA
AAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTG
CGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAGGGCCCAATACCGCGGAGAAAATGGCGAAAAACGTTGCG
ATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCCAAAAACTTCCTGGAAATGACAACAGCACGGC
AACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATAGTGAAAACAATCACGAATGACCAAATTGAAGTTACTAATGCTA
CTGAGCTGGTTCAGAATTCCTCAATAGGTGAAATATGCGACAGTCCTCATCAGATCCTTGATGGAGAAAACTGCACACTAATAGAT
GCTCTATTGGGAGACCCTCAGTGTGATGGCTTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGAAGCAAAGCCTACAGCAACTG
TTACCCTTATGATGTGCCGGATTATGCCTCCCTTAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTTTAACAATGAAAGCTTCA
ATTGGACTGGAGTCACTCAAAACGGAACAAGTTCTGCTTGCATAAGGAGATCTAATAATAGTTTCTTTAGTAGATTAAATTGGTTG
ACCCACTTAAACTTCAAATACCCAGCATTGAACGTGACTATGCCAAACAATGAACAATTTGACAAATTGTACATTTGGGGGGTTCA
CCACCCGGGTACGGACAAGGACCAAATCTTCCTGTATGCTCAATCATCAGGAAGAATCACAGTATCTACCAAAAGAAGCCAACAAG
CTGTAATCCCGAATATCGGATCTAGACCCAGAATAAGGAATATCCCTAGCAGAATAAGCATCTATTGGACAATAGTAAAACCGGGA
GACATACTTTTGATTAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAAATACGAAGTGGGAAAAGCTCAATAATGAG
ATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCAAATGGAAGCATTCCCAATGACAAACCATTCCAAAATGTAA
ACAGGATCACATACGGGGCCTGTCCCAGATATGTTAAGCAAAGCACTCTGAAATTGGCAACAGGAATGCGAAATGTACCAGAGAAA
CAAACTAGAGGCATATTTGGCGCAATAGCGGGTTTCATAGAAAATGGTTGGGAGGGAATGGTGGATGGTTGGTACGGTTTCAGGCA
TCAAAATTCTGAGGGAAGAGGACAAGCAGCAGATCTCAAAAGCACTCAAGCAGCAATCGATCAAATCAATGGGAAGCTGAATCGAT
TGATCGGGAAAACCAACGAGAAATTCCATCAGATTGAAAAAGAATTCTCAGAAGTCGAAGGGAGAATTCAGGACCTTGAGAAATAT
GTTGAGGACACTAAAATAGATCTCTGGTCATACAACGCGGAGCTTCTTGTTGCCCTGGAGAACCAACATACAATTGATCTAACTGA
CTCAGAAATGAACAAACTGTTTGAAAAAACAAAGAAGCAACTAAGGGAAAATGCTGAGGATATGGGCAATGGTTGTTTCAAAATAT
ACCACAAATGTGACAATGCCTGCATAGGATCAATCAGAAATGGAACTTATGACCACGATGTATACAGAGATGAAGCATTAAACAAC
CGGTTCCAGATCAAGGGAGTTGAGCTGAAGTCAGGGTACAAAGATTGGATCCTATGGATTTCCTTTGCCATATCATGTTTTTTGCT
TTGTGTTGCTTTGTTGGGGTTCATCATGTGGGCCTGCCAAAAGGGCAACATTAGGTGCAACATTTGCATTTGAAGGCCTATTTTCT
TTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAAT
TTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAAAAAAAAAAAAAA
AAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGT
TGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTA
TGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAA
TTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.44G:
Trình tự axit amin của PDISP/H3 Victoria (SEQ ID NO:122)

MAKNVAIFGLLFSLLVLVPSQIFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQNSSIGEICDSPHQILDGE
NCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFNNESFNWTGVTQNGTSSACIRRSNNSFF
SRLNWLTHLNFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKDQIFLYAQSSGRITVSTKRSQQAVIPNIGSRPRIRNIPSRISIYW
TIVKPGDILLINSTGNLIAPRGYFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRITYGACPRYVKQSTLKLATGM
RNVPEKQTRGIFGAIAGFIENGWEGMVDGWYGFRHQNSEGRGQAADLKSTQAAIDQINGKLNRLIGKTNEKFHQIEKEFSEVEGRI
QDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEKTKKQLRENAEDMGNGCFKIYHKCDNACIGSIRNGTYDHDVYR
DEALNNRFQIKGVELKSGYKDWILWISFAISCFLLCVALLGFIMWACQKGNIRCNICI*

Fig.44H
Sự biểu diễn theo sơ đồ của cấu trúc số 1800



2X3B-2X35S/CPMV-HT*/ PDISP/H3 Victoria/ NOS (Cấu trúc số 1819)

Fig.45A:
IF(SacII)-Kozac_PDI.c
GTCGGGCCCAATACCGCGGAGAAAATGGCGAAAAACGTTGCGATTTTC (SEQ ID NO.: 123)

Fig.45B:
IF-H3V36111.s1-4r
ACTAAAGAAAATAGGCCTTCAAATGCAAATGTTGCACCTAATGTTGCCCTT (SEQ ID NO: 124)

Fig.45C:
Sự biểu diễn theo sơ đồ của cấu trúc 2181. Vị trí enzym giới hạn SacII và StuI được sử dụng cho tuyến tính hoá plasmit được chú giải đối với sự biểu diễn.

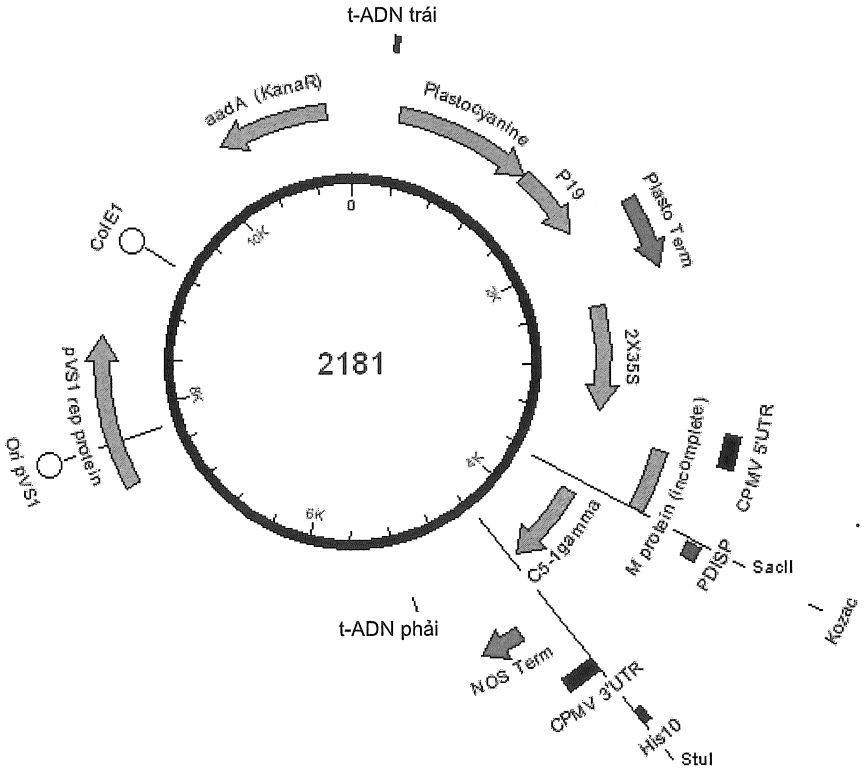


Fig.45D
Cấu trúc 2181 từ bên trái sang bên phải đường biên t-ADN (gạch dưới).
2X35S/CPMV-HT*/NOS với Plastoxyanin-P19-Plastoxyanin làm câm cát xét biểu hiện chất ức chế (SEQ ID NO:125)

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACTTAATAACACATTGCGGACGTTTTTAATGTACTGAATTAACG
CCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTTACT
TGAACAAAAATATTCACCTACTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTG
ATATTTTGACAACAATTTTGTTGCAACATTTGAGAAAATTTTGTTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAA
AGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCATT
GAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTAGAGA
ATTTTTGGCAAGTCATTAAAAAGAAAGAATAAATTATTTTTAAAATTAAAAGTTGAGTCATTTGATTAAACATGTGATTATTTAAT
GAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCTATAT

ATTGCCCCATAGAGTCAGTTAACTCATTTTTATATTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAAA
AAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATCCAATCCAACCAATCA
CAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCACACATC
TGAGCCACACAAAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAA
ACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGA
ACAAGCTAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCCTTCAAACTTCCTGACGAAAGTCCGAGTTGGACTG
AGTGGCGGCTACATAACGATGAGACGAATTCGAATCAAGATAATCCCCTTGGTTTCAAGGAAAGCTGGGGTTTCGGGAAAGTTGTA
TTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCACTGCACAGAGTCCTTGGATCTTGGACGGGAGATTCGGTTAACTATGC
AGCATCTCGATTTTTCGGTTTCGACCAGATCGGATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGT
CGCGAACTCTTCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGGAAAGTAAT
GTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAATGCTTCTTCGTCTCCTATTTATAAT
ATGGTTTGTTATTGTTAATTTTGTTCTTGTAGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTG
GTGTAATGTAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCCATAACTAACTAGACATGAAGACCTGCCGCGTACA
ATTGTCTTATATTTGAACAACTAAAATTGAACATCTTTTGCCACAACTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGT
TCAATAGATTAATAATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAACTACTAATTTTATATCATCCC
CTTTGATAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTGTCGTTTCCCGCCTTCAGTTTGCAAGCTGCTC
TAGCCGTGTAGCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGGAATTACTAGCGCGTGTCGACAAGCTTGCATGCCGGTCAACA
TGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAA
AGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTC
CTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCA
CGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACA
CTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAA
CCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATT
GCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAA
AAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTA
TCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAA
CGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGC
GATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACAC
GTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGT
TCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCT
GTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAA
CTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCAATACCGCGGAGAAAATGGCGAAAAACGTTGCG
ATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGACGTCACTCCTCAGCCAAAACGACACCCCCAT
CTGTCTATCCACTGGCCCCTGGATCTGCTGCCCAAACTAACTCCATGGTGACCCTGGGATGCCTGGTCAAGGGCTATTTCCCTGAG
CCAGTGACAGTGACCTGGAACTCTGGATCCCTGTCCAGCGGTGTGCACACCTTCCCAGCTGTCCTGCAGTCTGACCTCTACACTCT
GAGCAGCTCAGTGACTGTCCCCTCCAGCACCTGGCCCAGCGAGACCGTCACCTGCAACGTTGCCCACCCGGCCAGCAGCACCAAGG
TGGACAAGAAAATTGTGCCCAGGGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCATCTTCCCC
CCAAAGCCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGT
CCAGTTCAGCTGGTTTGTAGATGATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGCT
CAGTCAGTGAACTTCCCATCATGCACCAGGACTGGCTCAATGGCAAGGAGCGATCGCTCACCATCACCATCACCATCACCATCACC
ATTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGT
GTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTT
TATTAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTT
CTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGT
AATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAG
CGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCCCACGTGACT
AGTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTT
TCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCTAGAGCAGC
TTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAAACTATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGA
AAAGAGCGTTTA

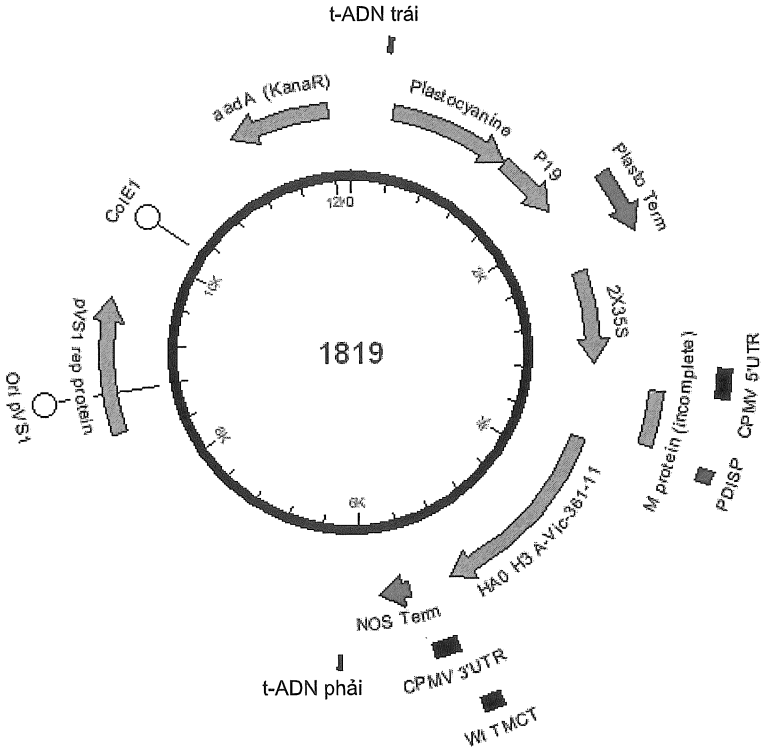
Fig.45E

Cát xét biểu hiện số 1819 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. Trình tự nucleotit gạch dưới. (SEQ ID NO:126)

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTT
TCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAG
GTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCC
CCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCA
CGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATAT
CCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGC
CATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCAT
CGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAAT

CCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAA
AAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTG
CGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTG
TGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGG
CTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATT
GTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGT
TTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCAATACCGCGGAGAAAATGGCGAAAAA
CGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCCAAAAACTTCCTGGAAATGACAACA
GCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATAGTGAAAACAATCACGAATGACCAAATTGAAGTTACT
AATGCTACTGAGCTGGTTCAGAATTCCTCAATAGGTGAAATATGCGACAGTCCTCATCAGATCCTTGATGGAGAAAACTGCACACT
AATAGATGCTCTATTGGGAGACCCTCAGTGTGATGGCTTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGAAGCAAAGCCTACA
GCAACTGTTACCCTTATGATGTGCCGGATTATGCCTCCCTTAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTTTAACAATGAA
AGCTTCAATTGGACTGGAGTCACTCAAAACGGAACAAGTTCTGCTTGCATAAGGAGATCTAATAATAGTTTCTTTAGTAGATTAAA
TTGGTTGACCCACTTAAACTTCAAATACCCAGCATTGAACGTGACTATGCCAAACAATGAACAATTTGACAAATTGTACATTTGGG
GGGTTCACCACCCGGGTACGGACAAGGACCAAATCTTCCTGTATGCTCAATCATCAGGAAGAATCACAGTATCTACCAAAAGAAGC
CAACAAGCTGTAATCCCGAATATCGGATCTAGACCCAGAATAAGGAATATCCCTAGCAGAATAAGCATCTATTGGACAATAGTAAA
ACCGGGAGACATACTTTTGATTAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAAATACGAAGTGGGAAAAGCTCAA
TAATGAGATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCAAATGGAAGCATTCCCAATGACAAACCATTCCAA
AATGTAAACAGGATCACATACGGGGCCTGTCCCAGATATGTTAAGCAAAGCACTCTGAAATTGGCAACAGGAATGCGAAATGTACC
AGAGAAACAAACTAGAGGCATATTTGGCGCAATAGCGGGTTTCATAGAAAATGGTTGGGAGGGAATGGTGGATGGTTGGTACGGTT
TCAGGCATCAAAATTCTGAGGGAAGAGGACAAGCAGCAGATCTCAAAAGCACTCAAGCAGCAATCGATCAAATCAATGGGAAGCTG
AATCGATTGATCGGGAAAACCAACGAGAAATTCCATCAGATTGAAAAAGAATTCTCAGAAGTCGAAGGGAGAATTCAGGACCTTGA
GAAATATGTTGAGGACACTAAAATAGATCTCTGGTCATACAACGCGGAGCTTCTTGTTGCCCTGGAGAACCAACATACAATTGATC
TAACTGACTCAGAAATGAACAAACTGTTTGAAAAAACAAAGAAGCAACTAAGGGAAAATGCTGAGGATATGGGCAATGGTTGTTTC
AAAATATACCACAAATGTGACAATGCCTGCATAGGATCAATCAGAAATGGAACTTATGACCACGATGTATACAGAGATGAAGCATT
AAACAACCGGTTCCAGATCAAGGGAGTTGAGCTGAAGTCAGGGTACAAAGATTGGATCCTATGGATTTCCTTTGCCATATCATGTT
TTTTGCTTTGTGTTGCTTTGTTGGGGTTCATCATGTGGGCCTGCCAAAAGGGCAACATTAGGTGCAACATTTGCATTTGAAGGCCT
ATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTT
ATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAAAAAAA
AAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGA
ATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACG
TTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTA
GGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.45F
Sự biểu diễn theo sơ đồ của cấu trúc số 1819



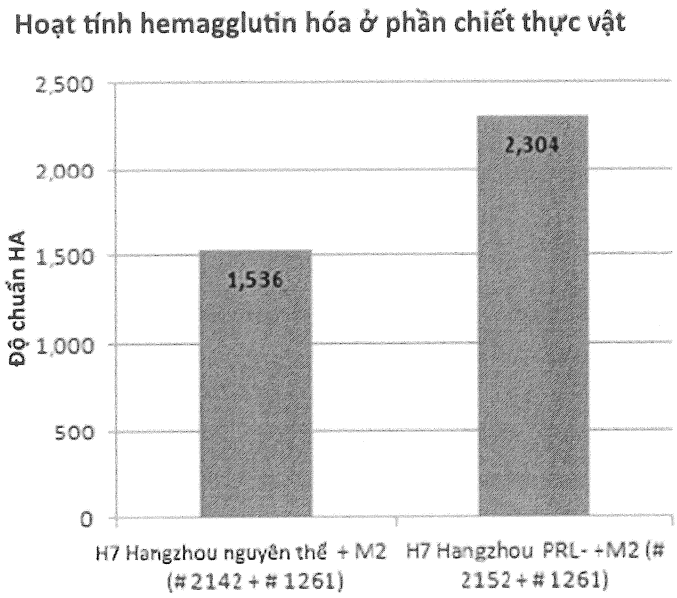


Fig. 46A

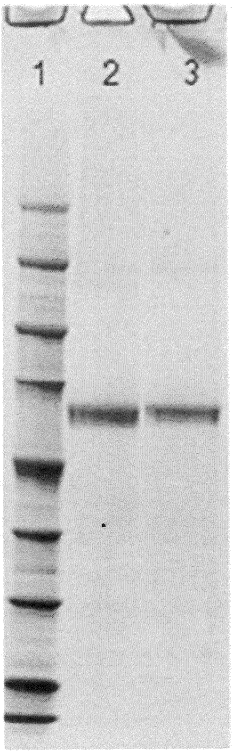


Fig. 46C

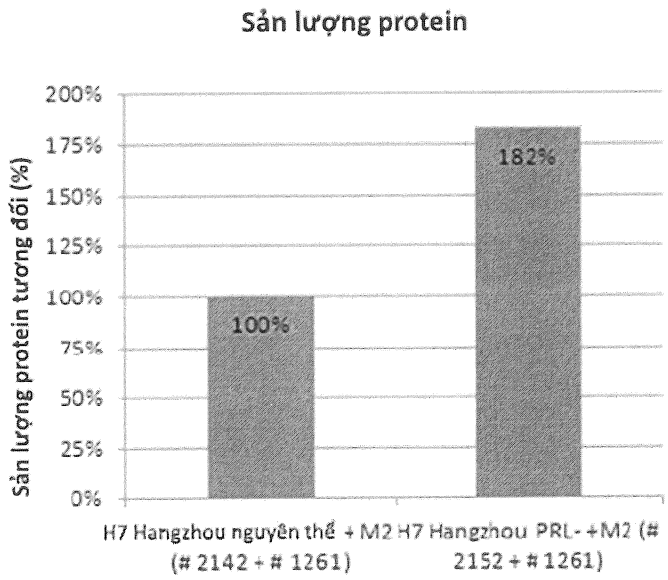


Fig.46B

Fig. 47A

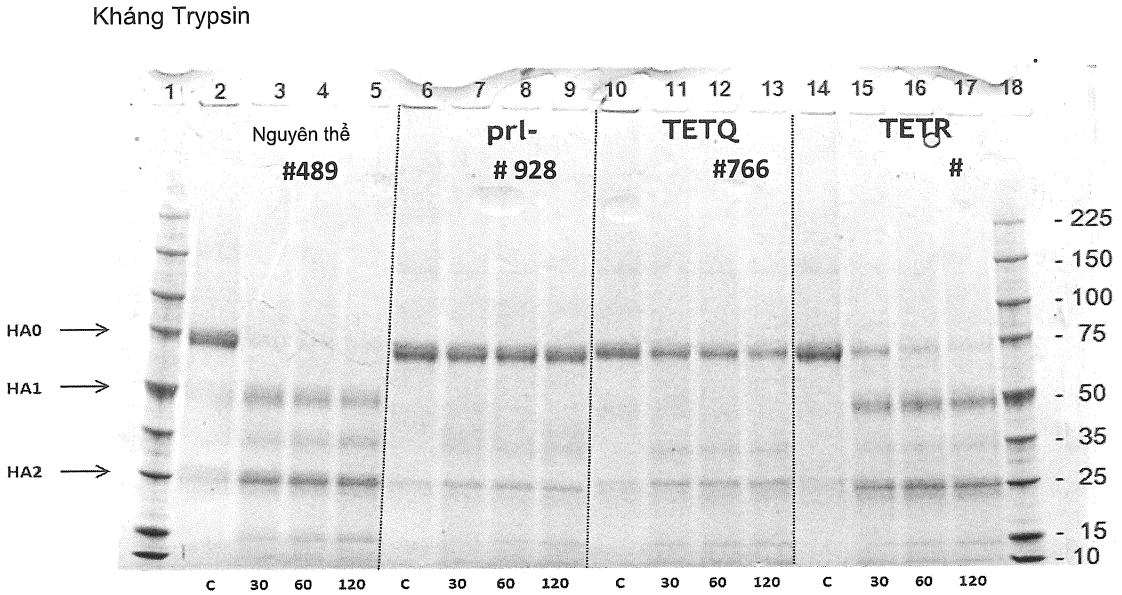
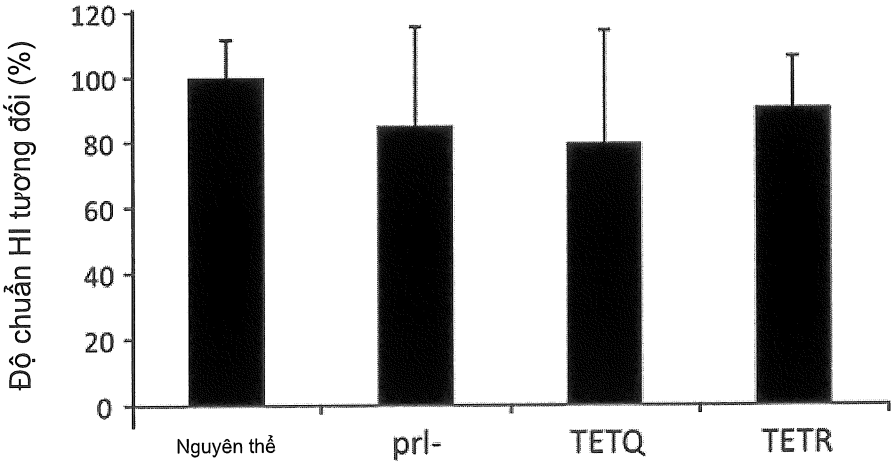


Fig. 47B



2X35S/CPMV HT+/ PDISP/H2 Singapore/ NOS (Cấu trúc số 2220)

Fig.48A: SEQ ID NO: 127

IF**-H2S157.s1-6r

ACTAAAGAAAATAGGCCTTCATATGCAGATCCTGCACTGCAGAGACCCGT

Fig.48B: SEQ ID NO: 128

Trình tự nucleotit của PDISP/H2 Singapore.

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGACCAAA
TATGCATTGGATACCATGCCAATAATTCCACAGAGAAGGTCGACACAATTCTAGAGCGGAACGTCACTGTGACTCATGC
CAAGGACATTCTTGAGAAGACCCATAACGGAAAGTTATGCAAACTAAACGGAATCCCTCCACTTGAACTAGGGGACTGT
AGCATTGCCGGATGGCTCCTTGGAAATCCAGAATGTGATAGGCTTCTAAGTGTGCCAGAATGGTCCTATATAATGGAGA
AAGAAAACCCGAGAGACGGTTTGTGTTATCCAGGCAGCTTCAATGATTATGAAGAATTGAAACATCTCCTCAGCAGCGT
GAAACATTTCGAGAAAGTAAAGATTCTGCCCAAAGATAGATGGACACAGCATACAACAACTGGAGGTTCACGGGCCTGC
GCGGTGTCTGGTAATCCATCATTCTTCAGGAACATGGTCTGGCTGACAAAGAAAGAATCAAATTATCCGGTTGCCAAAG
GATCGTACAACAATACAAGCGGAGAACAAATGCTAATAATTTGGGGGGTGCACCATCCCAATGATGAGACAGAACAAAG
AACATTGTACCAGAATGTGGGAACCTATGTTTCCGTAGGCACATCAACATTGAACAAAAGGTCAACCCCAGACATAGCA
ACAAGGCCTAAAGTGAATGGACTAGGAAGTAGAATGGAGTTCTCTTGGACCCTATTGGATATGTGGGACACCATAAATT
TTGAGAGTACTGGTAATCTAATTGCACCAGAGTATGGATTCAAAATATCGAAAAGAGGTAGTTCAGGGATCATGAAAAC
AGAAGGAACACTTGAGAACTGTGAGACCAAATGCCAAACTCCTTTGGGAGCAATAAATACAACATTGCCTTTTCACAAT
GTCCACCCACTGACAATAGGTGAGTGCCCCAAATATGTAAAATCGGAGAAGTTGGTCTTAGCAACAGGACTAAGGAATG
TTCCCCAGATTGAATCAAGAGGATTGTTTGGGGCAATAGCTGGTTTTATAGAAGGAGGATGGCAAGGAATGGTTGATGG
TTGGTATGGATACCATCACAGCAATGACCAGGGATCAGGGTATGCAGCAGACAAAGAATCCACTCAAAAGGCATTTGAT
GGAATCACCAACAAGGTAAATTCTGTGATTGAAAAGATGAACACCCAATTTGAAGCTGTTGGGAAAGAGTTCAGTAACT
TAGAGAGAAGACTGGAGAACTTGAACAAAAAGATGGAAGACGGGTTTCTAGATGTGTGGACATACAATGCTGAGCTTCT
AGTTCTGATGGAAAATGAGAGGACACTTGACTTTCATGATTCTAATGTCAAGAATCTGTATGATAAAGTCAGAATGCAG
CTGAGAGACAACGTCAAAGAACTAGGAAATGGATGTTTTGAATTTTATCACAAATGTGATGATGAATGCATGAATAGTG
TGAAAAACGGGACGTATGATTATCCCAAGTATGAAGAAGAGTCTAAACTAAATAGAAATGAAATCAAAGGGGTAAAATT
GAGCAGCATGGGGGGTTTATCAAATCCTTGCCATTTATGCTACAGTAGCAGGTTCTCTGTCACTGGCAATCATGATGGCT
GGGATCTCTTTCTGGATGTGCTCCAACGGGTCTCTGCAGTGCAGGATCTGCATATGA

Fig.48C: SEQ ID NO:129
Cát xét biểu hiện số 2220 từ đoạn khởi đầu 2X35S đến đoạn kết thúc. Trình tự nucleotit của PDISP/H5 Singapore được gạch dưới.

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTG
AGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGAT
AGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGAC
AGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAG
TGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCA
AAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTT
ATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATG
CCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTC
TTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCT
TCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAA
ACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGA
TCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTG
GACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCT
GGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCT
TGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGA

AACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCC
AATACCGCGGAGAAAATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGA
TCTTCGCGGACCAAATATGCATTGGATACCATGCCAATAATTCCACAGAGAAGGTCGACACAATTCTAGAGCGGAACGT
CACTGTGACTCATGCCAAGGACATTCTTGAGAAGACCCATAACGGAAAGTTATGCAAACTAAACGGAATCCCTCCACTT
GAACTAGGGGACTGTAGCATTGCCGGATGGCTCCTTGGAAATCCAGAATGTGATAGGCTTCTAAGTGTGCCAGAATGGT
CCTATATAATGGAGAAAGAAAACCCGAGAGACGGTTTGTGTTATCCAGGCAGCTTCAATGATTATGAAGAATTGAAACA
TCTCCTCAGCAGCGTGAAACATTTCGAGAAAGTAAAGATTCTGCCCAAAGATAGATGGACACAGCATACAACAACTGGA
GGTTCACGGGCCTGCGCGGTGTCTGGTAATCCATCATTCTTCAGGAACATGGTCTGGCTGACAAAGAAAGAATCAAATT
ATCCGGTTGCCAAAGGATCGTACAACAATACAAGCGGAGAACAAATGCTAATAATTTGGGGGGTGCACCATCCCAATGA
TGAGACAGAACAAAGAACATTGTACCAGAATGTGGGAACCTATGTTTCCGTAGGCACATCAACATTGAACAAAAGGTCA
ACCCCAGACATAGCAACAAGGCCTAAAGTGAATGGACTAGGAAGTAGAATGGAGTTCTCTTGGACCCTATTGGATATGT
GGGACACCATAAATTTTGAGAGTACTGGTAATCTAATTGCACCAGAGTATGGATTCAAAATATCGAAAAGAGGTAGTTC
AGGGATCATGAAAACAGAAGGAACACTTGAGAACTGTGAGACCAAATGCCAAACTCCTTTGGGAGCAATAAATACAACA
TTGCCTTTTCACAATGTCCACCCACTGACAATAGGTGAGTGCCCCAAATATGTAAAATCGGAGAAGTTGGTCTTAGCAA
CAGGACTAAGGAATGTTCCCCAGATTGAATCAAGAGGATTGTTTGGGGCAATAGCTGGTTTTATAGAAGGAGGATGGCA
AGGAATGGTTGATGGTTGGTATGGATACCATCACAGCAATGACCAGGGATCAGGGTATGCAGCAGACAAAGAATCCACT
CAAAAGGCATTTGATGGAATCACCAACAAGGTAAATTCTGTGATTGAAAAGATGAACACCCAATTTGAAGCTGTTGGGA
AAGAGTTCAGTAACTTAGAGAGAAGACTGGAGAACTTGAACAAAAAGATGGAAGACGGGTTTCTAGATGTGTGGACATA
CAATGCTGAGCTTCTAGTTCTGATGGAAAATGAGAGGACACTTGACTTTCATGATTCTAATGTCAAGAATCTGTATGAT
AAAGTCAGAATGCAGCTGAGAGACAACGTCAAAGAACTAGGAAATGGATGTTTTGAATTTTATCACAAATGTGATGATG
AATGCATGAATAGTGTGAAAAACGGGACGTATGATTATCCCAAGTATGAAGAAGAGTCTAAACTAAATAGAAATGAAAT
CAAAGGGGTAAAATTGAGCAGCATGGGGGTTTATCAAATCCTTGCCATTTATGCTACAGTAGCAGGTTCTCTGTCACTG
GCAATCATGATGGCTGGGATCTCTTTCTGGATGTGCTCCAACGGGTCTCTGCAGTGCAGGATCTGCATATGAAGGCCTA
TTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTT
TATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAT
TTTATTAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCA
ATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCAT
GTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAAT
ACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

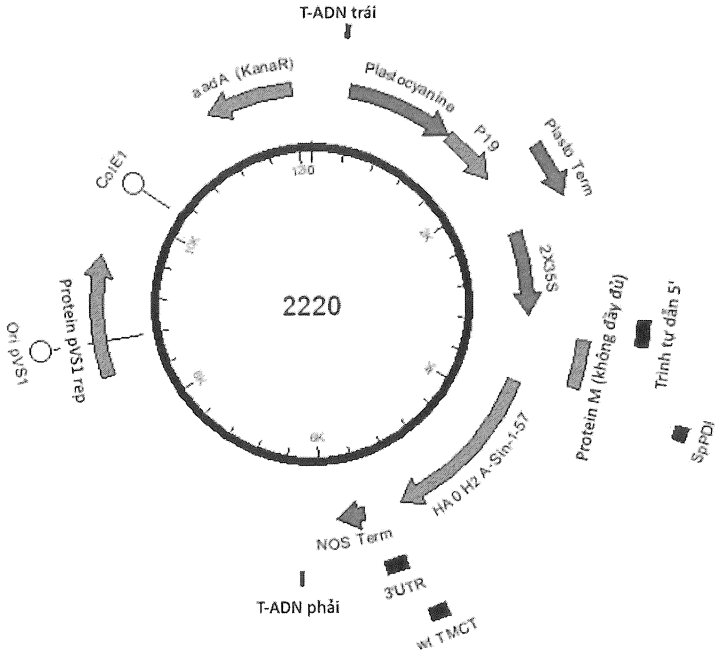
Fig.48D: SEQ ID NO:130

Trình tự axit amin của PDISP/H2 Singapore

MAKNVAIFGLLFSLLVLVPSQIFADQICIGYHANNSTEKVDTILERNVTVTHAKDILEKTHNGKLCKLNGIPPLELGDC
SIAGWLLGNPECDRLLSVPEWSYIMEKENPRDGLCYPGSFNDYEELKHLLSSVKHFEKVKILPKDRWTQHTTTGGSRAC
AVSGNPSFFRNMVWLTKKESNYPVAKGSYNNTSGEQMLIIWGVHHPNDETEQRTLYQNVGTYVSVGTSTLNKRSTPDIA
TRPKVNGLGSRMEFSWTLLDMWDTINFESTGNLIAPEYGFKISKRGSSGIMKTEGTLENCETKCQTPLGAINTTLPFHN
VHPLTIGECPKYVKSEKLVLATGLRNVPQIESRGLFGAIAGFIEGGWQGMVDGWYGYHHSNDQGSGYAADKESTQKAFD
GITNKVNSVIEKMNTQFEAVGKEFSNLERRLENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRMQ
LRDNVKELGNGCFEFYHKCDDECMNSVKNGTYDYPKYEEESKLNRNEIKGVKLSSMGVYQILAIYATVAGSLSLAIMMA
GISFWMCSNGSLQCRICI*

Fig. 48E

Biểu diễn theo sơ đồ của cấu trúc 2220



2X35S/CPMV HT+/ PDISP/H2 Singapore với vòng phân giải protein loại bớt/NOS (Cấu trúc số 2221)

Fig.49A: SEQ ID NO: 131

H2S157(Pri-).r

TGCCATCCTCCGCCGGGAACATTCCTTAGTCCTGTTGCTAAGACCAAC

Fig.49B: SEQ ID NO: 132

H2S157(Pri-).c

AGGAATGTTCCCGGCGGAGGATGGCAAGGAATGGTTGATGGTTGGTATGG

Fig.49C: SEQ ID NO: 133

Cát xét biểu hiện số 2221 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. PDISP/H2 Singapore với trình tự nucleotit vòng phân giải protein loại bớt được gạch dưới

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTG
AGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGAT
AGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGAC
AGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAG
TGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCA
AAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTT
ATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATG
CCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTC
TTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCT
TCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAA
ACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGA

TCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTG
GACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCT
GGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCT
TGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGA
AACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCC
AATACCGCGGAGAAAATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGA
TCTTCGCGGACCAAATATGCATTGGATACCATGCCAATAATTCCACAGAGAAGGTCGACACAATTCTAGAGCGGAACGT
CACTGTGACTCATGCCAAGGACATTCTTGAGAAGACCCATAACGGAAAGTTATGCAAACTAAACGGAATCCCTCCACTT
GAACTAGGGGACTGTAGCATTGCCGGATGGCTCCTTGGAAATCCAGAATGTGATAGGCTTCTAAGTGTGCCAGAATGGT
CCTATATAATGGAGAAAGAAAACCCGAGAGACGGTTTGTGTTATCCAGGCAGCTTCAATGATTATGAAGAATTGAAACA
TCTCCTCAGCAGCGTGAAACATTTCGAGAAAGTAAAGATTCTGCCCAAAGATAGATGGACACAGCATACAACAACTGGA
GGTTCACGGGCCTGCGCGGTGTCTGGTAATCCATCATTCTTCAGGAACATGGTCTGGCTGACAAAGAAAGAATCAAATT
ATCCGGTTGCCAAAGGATCGTACAACAATACAAGCGGAGAACAAATGCTAATAATTTGGGGGGTGCACCATCCCAATGA
TGAGACAGAACAAAGAACATTGTACCAGAATGTGGGAACCTATGTTTCCGTAGGCACATCAACATTGAACAAAAGGTCA
ACCCCAGACATAGCAACAAGGCCTAAAGTGAATGGACTAGGAAGTAGAATGGAGTTCTCTTGGACCCTATTGGATATGT
GGGACACCATAAATTTTGAGAGTACTGGTAATCTAATTGCACCAGAGTATGGATTCAAAATATCGAAAAGAGGTAGTTC
AGGGATCATGAAAACAGAAGGAACACTTGAGAACTGTGAGACCAAATGCCAAACTCCTTTGGGAGCAATAAATACAACA
TTGCCTTTTCACAATGTCCACCCACTGACAATAGGTGAGTGCCCCAAATATGTAAAATCGGAGAAGTTGGTCTTAGCAA
CAGGACTAAGGAATGTTCCCGGCGGAGGATGGCAAGGAATGGTTGATGGTTGGTATGGATACCATCACAGCAATGACCA
GGGATCAGGGTATGCAGCAGACAAAGAATCCACTCAAAAGGCATTTGATGGAATCACCAACAAGGTAAATTCTGTGATT
GAAAAGATGAACACCCAATTTGAAGCTGTTGGGAAAGAGTTCAGTAACTTAGAGAGAAGACTGGAGAACTTGAACAAAA
AGATGGAAGACGGGTTTCTAGATGTGTGGACATACAATGCTGAGCTTCTAGTTCTGATGGAAAATGAGAGGACACTTGA
CTTTCATGATTCTAATGTCAAGAATCTGTATGATAAAGTCAGAATGCAGCTGAGAGACAACGTCAAAGAACTAGGAAAT
GGATGTTTTGAATTTTATCACAAATGTGATGATGAATGCATGAATAGTGTGAAAAACGGGACGTATGATTATCCCAAGT
ATGAAGAAGAGTCTAAACTAAATAGAAATGAAATCAAAGGGGTAAAATTGAGCAGCATGGGGGTTTATCAAATCCTTGC
CATTTATGCTACAGTAGCAGGTTCTCTGTCACTGGCAATCATGATGGCTGGGATCTCTTTCTGGATGTGCTCCAACGGG
TCTCTGCAGTGCAGGATCTGCATATGAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATG
TTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGT
CGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAG
CTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGAT
TATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTT
TTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTAT
CGCGCGCGGTGTCATCTATGTTACTAGAT

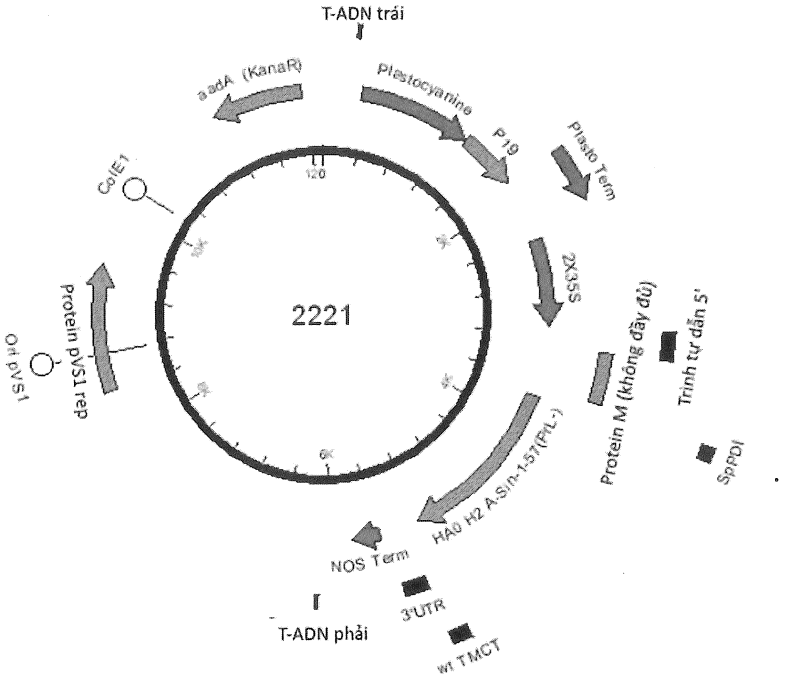
Fig.49D: SEQ ID NO:134

Trình tự axit amin của PDISP/H2 Singapore với vòng phân giải protein loại bớt

MAKNVAIFGLLFSLLVLVPSQIFADQICIGYHANNSTEKVDTILERNVTVTHAKDILEKTHNGKLCKLNGIPPLELGDCSIAGWLLGNPECD
RLLSVPEWSYIMEKENPRDGLCYPGSFNDYEELKHLLSSVKHFEKVKILPKDRWTQHTTTGGSRACAVSGNPSFFRNMVWLTKKESNYP
VAKGSYNNTSGEQMLIIWGVHHPNDETEQRTLYQNVGTYVSVGTSTLNKRSTPDIATRPKVNGLGSRMEFSWTLLDMWDTINFESTG
NLIAPEYGFKISKRGSSGIMKTEGTLENCETKCQTPLGAINTTLPFHNVHPLTIGECPKYVKSEKLVLATGLRNVPGGGWQGMVDGWYG
YHHSNDQGSGYAADKESTQKAFDGITNKVNSVIEKMNTQFEAVGKEFSNLERRLENLNKKMEDGFLDVWTYNAELLVLMENERTLDF
HDSNVKNLYDKVRMQLRDNVKELGNGCFEFYHKCDDECMNSVKNGTYDYPKYEEESKLNRNEIKGVKLSSMGVYQILAIYATVAGSLSL
AIMMAGISFWMCSNGSLQCRICI*

Fig. 49E

Biểu diễn theo sơ đồ của cấu trúc số 2221



PDISP/H2 Singapore (Cấu trúc số 2222) và PDISP/H2 Singapore với vòng phân giải protein loại bót (Cấu trúc số 2223) trong hệ biểu hiện 2X35S/CPMV 160+/NOS

Fig.50A: SEQ ID NO:135

Cát xét biểu hiện số 2222 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. Trình tự nucleotit PDISP/H2 Singapore được gạch dưới

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTG
AGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGAT
AGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGAC
AGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAG
TGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCA
AAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTT
ATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATG
CCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTC
TTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCT
TCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAA
ACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGA
TCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAGGGCCCAATACCGCGGAGAAAATGGCGAAAAACGTTGCGAT
TTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGACCAAATATGCATTGGATACCATGCC
AATAATTCCACAGAGAAGGTCGACACAATTCTAGAGCGGAACGTCACTGTGACTCATGCCAAGGACATTCTTGAGAAGA
CCCATAACGGAAAGTTATGCAAACTAAACGGAATCCCTCCACTTGAACTAGGGGACTGTAGCATTGCCGGATGGCTCCT
TGGAAATCCAGAATGTGATAGGCTTCTAAGTGTGCCAGAATGGTCCTATATAATGGAGAAAGAAAACCCGAGAGACGGT
TTGTGTTATCCAGGCAGCTTCAATGATTATGAAGAATTGAAACATCTCCTCAGCAGCGTGAAACATTTCGAGAAAGTAA
AGATTCTGCCCAAAGATAGATGGACACAGCATACAACAACTGGAGGTTCACGGGCCTGCGCGGTGTCTGGTAATCCATC
ATTCTTCAGGAACATGGTCTGGCTGACAAAGAAAGAATCAAATTATCCGGTTGCCAAAGGATCGTACAACAATACAAGC
GGAGAACAAATGCTAATAATTTGGGGGGTGCACCATCCCAATGATGAGACAGAACAAAGAACATTGTACCAGAATGTGG
GAACCTATGTTTCCGTAGGCACATCAACATTGAACAAAAGGTCAACCCCAGACATAGCAACAAGGCCTAAAGTGAATGG
ACTAGGAAGTAGAATGGAGTTCTCTTGGACCCTATTGGATATGTGGGACACCATAAATTTTGAGAGTACTGGTAATCTA

ATTGCACCAGAGTATGGATTCAAAATATCGAAAAGAGGTAGTTCAGGGATCATGAAAACAGAAGGAACACTTGAGAACT
GTGAGACCAAATGCCAAACTCCTTTGGGAGCAATAAATACAACATTGCCTTTTCACAATGTCCACCCACTGACAATAGG
TGAGTGCCCCAAATATGTAAAATCGGAGAAGTTGGTCTTAGCAACAGGACTAAGGAATGTTCCCCAGATTGAATCAAGA
GGATTGTTTGGGGCAATAGCTGGTTTTATAGAAGGAGGATGGCAAGGAATGGTTGATGGTTGGTATGGATACCATCACA
GCAATGACCAGGGATCAGGGTATGCAGCAGACAAAGAATCCACTCAAAAGGCATTTGATGGAATCACCAACAAGGTAAA
TTCTGTGATTGAAAAGATGAACACCCAATTTGAAGCTGTTGGGAAAGAGTTCAGTAACTTAGAGAGAAGACTGGAGAAC
TTGAACAAAAAGATGGAAGACGGGTTTCTAGATGTGTGGACATACAATGCTGAGCTTCTAGTTCTGATGGAAAATGAGA
GGACACTTGACTTTCATGATTCTAATGTCAAGAATCTGTATGATAAAGTCAGAATGCAGCTGAGAGACAACGTCAAAGA
ACTAGGAAATGGATGTTTTGAATTTTATCACAAATGTGATGATGAATGCATGAATAGTGTGAAAAACGGGACGTATGAT
TATCCCAAGTATGAAGAAGAGTCTAAACTAAATAGAAATGAAATCAAAGGGGTAAAATTGAGCAGCATGGGGGTTTATC
AAATCCTTGCCATTTATGCTACAGTAGCAGGTTCTCTGTCACTGGCAATCATGATGGCTGGGATCTCTTTCTGGATGTG
CTCCAACGGGTCTCTGCAGTGCAGGATCTGCATATGAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTG
CATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTG
TTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAGACCGGGAATT
CGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTC
TTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTAT
GAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAG
GATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig.50B: SEQ ID NO:136

Cát xét biểu hiện số 2223 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. PDISP/H2 Singapore với trình tự nucleotit vòng phân giải protein được gạch dưới.

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTG
AGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGAT
AGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGAC
AGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAG
TGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCA
AAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTT
ATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATG
CCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTC
TTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCT
TCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAA
ACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGA
TCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAGGGCCCAATACCGCGGAGAAAATGGCGAAAAACGTTGCGAT
TTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGACCAAATATGCATTGGATACCATGCC
AATAATTCCACAGAGAAGGTCGACACAATTCTAGAGCGGAACGTCACTGTGACTCATGCCAAGGACATTCTTGAGAAGA
CCCATAACGGAAAGTTATGCAAACTAAACGGAATCCCTCCACTTGAACTAGGGGACTGTAGCATTGCCGGATGGCTCCT
TGGAAATCCAGAATGTGATAGGCTTCTAAGTGTGCCAGAATGGTCCTATATAATGGAGAAAGAAAACCCGAGAGACGGT
TTGTGTTATCCAGGCAGCTTCAATGATTATGAAGAATTGAAACATCTCCTCAGCAGCGTGAAACATTTCGAGAAAGTAA
AGATTCTGCCCAAAGATAGATGGACACAGCATACAACAACTGGAGGTTCACGGGCCTGCGCGGTGTCTGGTAATCCATC
ATTCTTCAGGAACATGGTCTGGCTGACAAAGAAAGAATCAAATTATCCGGTTGCCAAAGGATCGTACAACAATACAAGC
GGAGAACAAATGCTAATAATTTGGGGGGTGCACCATCCCAATGATGAGACAGAACAAAGAACATTGTACCAGAATGTGG
GAACCTATGTTTCCGTAGGCACATCAACATTGAACAAAAGGTCAACCCCAGACATAGCAACAAGGCCTAAAGTGAATGG
ACTAGGAAGTAGAATGGAGTTCTCTTGGACCCTATTGGATATGTGGGACACCATAAATTTTGAGAGTACTGGTAATCTA
ATTGCACCAGAGTATGGATTCAAAATATCGAAAAGAGGTAGTTCAGGGATCATGAAAACAGAAGGAACACTTGAGAACT
GTGAGACCAAATGCCAAACTCCTTTGGGAGCAATAAATACAACATTGCCTTTTCACAATGTCCACCCACTGACAATAGG
TGAGTGCCCCAAATATGTAAAATCGGAGAAGTTGGTCTTAGCAACAGGACTAAGGAATGTTCCCGGCGGAGGATGGCAA
GGAATGGTTGATGGTTGGTATGGATACCATCACAGCAATGACCAGGGATCAGGGTATGCAGCAGACAAAGAATCCACTC
AAAAGGCATTTGATGGAATCACCAACAAGGTAAATTCTGTGATTGAAAAGATGAACACCCAATTTGAAGCTGTTGGGAA
AGAGTTCAGTAACTTAGAGAGAAGACTGGAGAACTTGAACAAAAAGATGGAAGACGGGTTTCTAGATGTGTGGACATAC
AATGCTGAGCTTCTAGTTCTGATGGAAAATGAGAGGACACTTGACTTTCATGATTCTAATGTCAAGAATCTGTATGATA
AAGTCAGAATGCAGCTGAGAGACAACGTCAAAGAACTAGGAAATGGATGTTTTGAATTTTATCACAAATGTGATGATGA
ATGCATGAATAGTGTGAAAAACGGGACGTATGATTATCCCAAGTATGAAGAAGAGTCTAAACTAAATAGAAATGAAATC

AAAGGGGTAAAATTGAGCAGCATGGGGGTTTATCAAATCCTTGCCATTTATGCTACAGTAGCAGGTTCTCTGTCACTGG
CAATCATGATGGCTGGGATCTCTTTCTGGATGTGCTCCAACGGGTCTCTGCAGTGCAGGATCTGCATATGAAGGCCTAT
TTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTT
ATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATT
TTATTAAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAA
TAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATG
TAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATA
CGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig. 50C

Biểu diễn bằng sơ đồ của cấu trúc số 2222

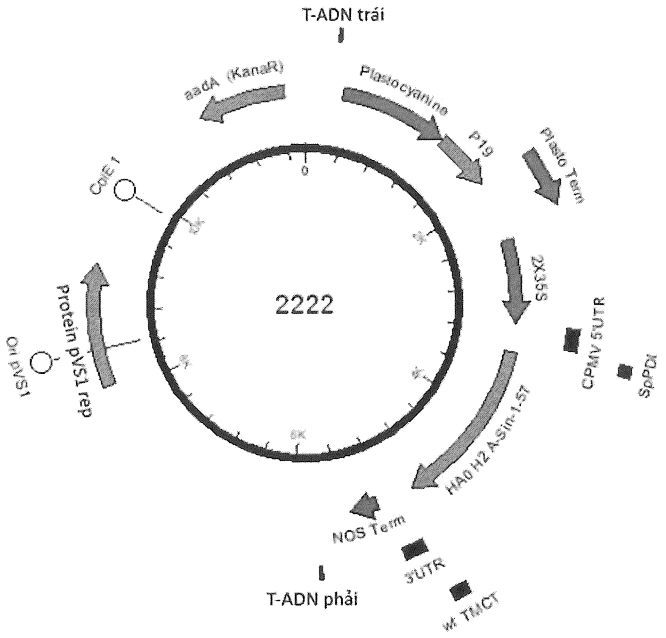
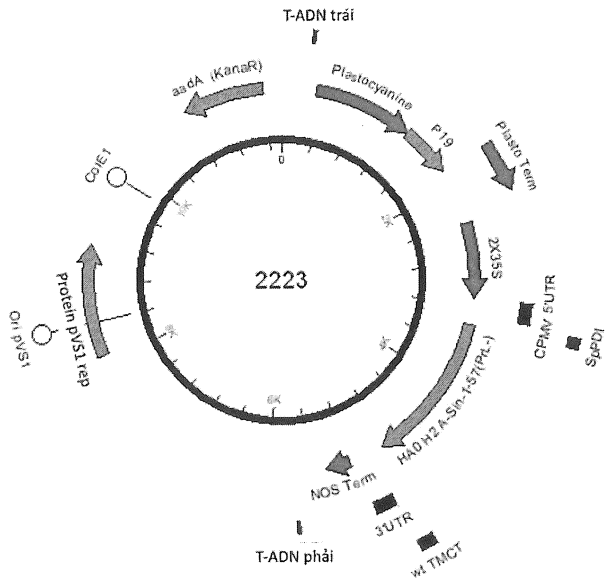


Fig. 50D

Biểu diễn theo sơ đồ của cấu trúc số 2223



2X35S/CPMV HT+ (cấu trúc số 2019) và 160+ (cấu trúc số 2139) đối với PDISP/H3 Perth

Fig.51A: SEQ ID NO:137

Trình tự nucleotit của PDISP/H3 Perth

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGCAAAAAC
TTCCTGGAAATGACAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATAGTGAAAACAAT
CACGAATGACCAAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAGTTCCTCAACAGGTGAAATATGCGACAGTCCT
CATCAGATCCTTGATGGAAAAAACTGCACACTAATAGATGCTCTATTGGGAGACCCTCAGTGTGATGGCTTCCAAAATA
AGAAATGGGACCTTTTTGTTGAACGCAGCAAAGCCTACAGCAACTGTTACCCTTATGATGTGCCGGATTATGCCTCCCT
TAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTTTAACAATGAAAGCTTCAATTGGACTGGAGTCACTCAAAACGGA
ACAAGCTCTGCTTGCATAAGGAGATCTAAAAACAGTTTCTTTAGTAGATTGAATTGGTTGACCCACTTAAACTTCAAAT
ACCCAGCATTGAACGTGACTATGCCAAACAATGAACAATTTGACAAATTGTACATTTGGGGGGTTCACCACCCGGGTAC
GGACAAAGACCAAATCTTCCTGTATGCTCAAGCATCAGGAAGAATCACAGTCTCTACCAAAAGAAGCCAACAAACCGTA
AGCCCGAATATCGGATCTAGACCCAGAGTAAGGAATATCCCTAGCAGAATAAGCATCTATTGGACAATAGTAAAACCGG
GAGACATACTTTTGATTAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAAATACGAAGTGGGAAAAGCTC
AATAATGAGATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCAAATGGAAGCATTCCCAATGACAAA
CCATTCCAAAATGTAAACAGGATCACATACGGGGCCTGTCCCAGATATGTTAAGCAAAACACTCTGAAATTGGCAACAG
GGATGCGAAATGTACCAGAGAAACAAACTAGAGGCATATTTGGCGCAATCGCGGGTTTCATAGAAAATGGTTGGGAGGG
AATGGTGGATGGTTGGTACGGTTTCAGGCATCAAAATTCTGAGGGAAGAGGACAAGCAGCAGATCTCAAAAGCACTCAA
GCAGCAATCGATCAAATCAATGGGAAGCTGAATAGATTGATCGGGAAAACCAACGAGAAATTCCATCAGATTGAAAAAG
AATTCTCAGAAGTCGAAGGGAGAATTCAGGACCTTGAGAAATATGTTGAGGACACTAAAATAGATCTCTGGTCATACAA
CGCGGAGCTTCTTGTTGCCCTGGAGAACCAACATACAATTGATCTAACTGACTCAGAAATGAACAAACTGTTTGAAAAA
ACAAAGAAGCAACTGAGGGAAAATGCTGAGGATATGGGCAATGGTTGTTTCAAAATATACCACAAATGTGACAATGCCT
GCATAGGATCAATCAGAAATGGAACTTATGACCACGATGTATACAGAGATGAAGCATTAAACAACCGGTTTCAGATCAA
GGGAGTTGAGCTGAAGTCAGGGTACAAAGATTGGATCCTATGGATTTCCTTTGCCATATCATGTTTTTTGCTTTGTGTT
GCTTTGTTGGGGTTCATCATGTGGGCCTGCCAAAAAGGCAACATTAGGTGCAACATTTGCATTTGA

Fig.51B: SEQ ID NO: 138

IF**-H3P1609.S1-6r

ACTAAAGAAAATAGGCCTTCAAATGCAAATGTTGCACCTAATGTTGCCTT

Fig.51C: . SEQ ID NO: 139

Trình tự axit amin của PDISP/H3 Perth

MAKNVAIFGLLFSLLVLVPSQIFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGEICDSP
HQILDGKNCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFNNESFNWTGVTQNG
TSSACIRRSKNSFFSRLNWLTHLNFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKDQIFLYAQASGRITVSTKRSQQTV
SPNIGSRPRVRNIPSRISIYWTIVKPGDILLINSTGNLIAPRGYFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDK
PFQNVNRITYGACPRYVKQNTLKLATGMRNVPEKQTRGIFGAIAGFIENGWEGMVDGWYGFRHQNSEGRGQAADLKSTQ
AAIDQINGKLNRLIGKTNEKFHQIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEK
TKKQLRENAEDMGNGCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWISFAISCFLLCV
ALLGFIMWACQKGNIRCNICI*

Fig. 51D

Biểu diễn bằng sơ đồ của cấu trúc số 2019

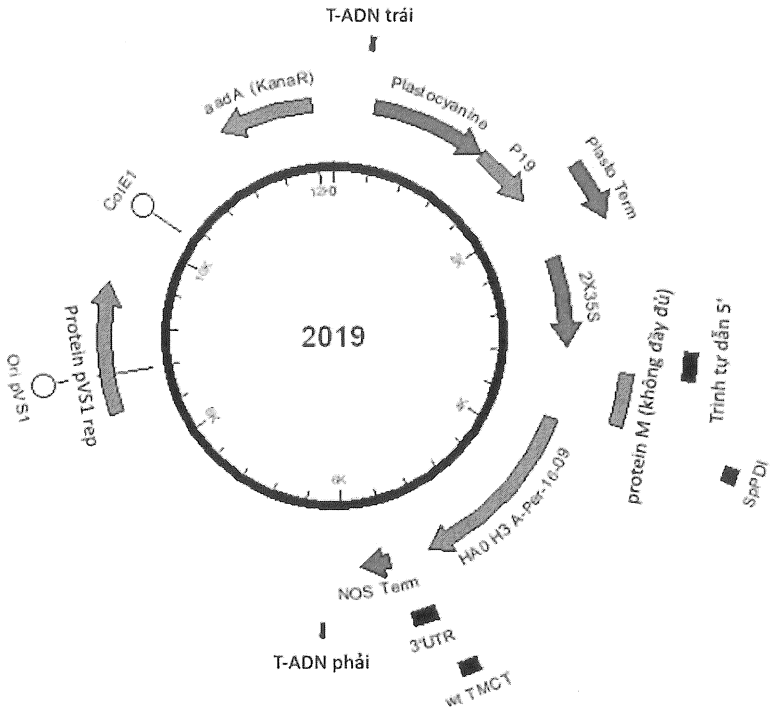
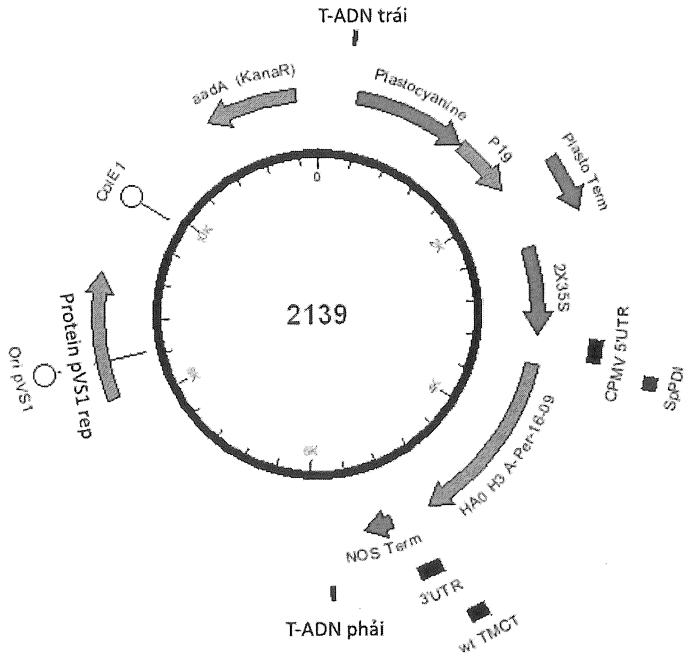


Fig. 51E

Biểu diễn bằng sơ đồ của cấu trúc số 2139



2X35S/CPMV HT+ (cấu trúc số 2039) và 160+ (cấu trúc số 2159) đối với PDISP/H3 Perth với vòng phân giải protein loại bớt

Fig.52A: SEQ ID NO:140

Trình tự nucleotit của PDISP/H3 Perth với vòng phân giải protein loại bớt

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGCAAAAAC
TTCCTGGAAATGACAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATAGTGAAAACAAT
CACGAATGACCAAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAGTTCCTCAACAGGTGAAATATGCGACAGTCCT
CATCAGATCCTTGATGGAAAAAACTGCACACTAATAGATGCTCTATTGGGAGACCCTCAGTGTGATGGCTTCCAAAATA
AGAAATGGGACCTTTTTGTTGAACGCAGCAAAGCCTACAGCAACTGTTACCCTTATGATGTGCCGGATTATGCCTCCCT
TAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTTTAACAATGAAAGCTTCAATTGGACTGGAGTCACTCAAAACGGA
ACAAGCTCTGCTTGCATAAGGAGATCTAAAAACAGTTTCTTTAGTAGATTGAATTGGTTGACCCACTTAAACTTCAAAT
ACCCAGCATTGAACGTGACTATGCCAAACAATGAACAATTTGACAAATTGTACATTTGGGGGGTTCACCACCCGGGTAC
GGACAAAGACCAAATCTTCCTGTATGCTCAAGCATCAGGAAGAATCACAGTCTCTACCAAAAGAAGCCAACAAACCGTA
AGCCCGAATATCGGATCTAGACCCAGAGTAAGGAATATCCCTAGCAGAATAAGCATCTATTGGACAATAGTAAAACCGG
GAGACATACTTTTGATTAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAAATACGAAGTGGGAAAAGCTC
AATAATGAGATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCAAATGGAAGCATTCCCAATGACAAA
CCATTCCAAAATGTAAACAGGATCACATACGGGGCCTGTCCCAGATATGTTAAGCAAAACACTCTGAAATTGGCAACAG
GGATGCGAAATGTACCAGGCGGAGGTTGGGAGGGAATGGTGGATGGTTGGTACGGTTTCAGGCATCAAAATTCTGAGGG
AAGAGGACAAGCAGCAGATCTCAAAAGCACTCAAGCAGCAATCGATCAAATCAATGGGAAGCTGAATAGATTGATCGGG
AAAACCAACGAGAAATTCCATCAGATTGAAAAAGAATTCTCAGAAGTCGAAGGGAGAATTCAGGACCTTGAGAAATATG
TTGAGGACACTAAAATAGATCTCTGGTCATACAACGCGGAGCTTCTTGTTGCCCTGGAGAACCAACATACAATTGATCT
AACTGACTCAGAAATGAACAAACTGTTTGAAAAAACAAAGAAGCAACTGAGGGAAAATGCTGAGGATATGGGCAATGGT
TGTTTCAAAATATACCACAAATGTGACAATGCCTGCATAGGATCAATCAGAAATGGAACTTATGACCACGATGTATACA
GAGATGAAGCATTAAACAACCGGTTTCAGATCAAGGGAGTTGAGCTGAAGTCAGGGTACAAAGATTGGATCCTATGGAT
TTCCTTTGCCATATCATGTTTTTTGCTTTGTGTTGCTTTGTTGGGGTTCATCATGTGGGCCTGCCAAAAAGGCAACATT
AGGTGCAACATTTGCATTTGA

Fig.52B: SEQ ID NO: 141

H3P1609(Prl-)#2.r

CCCTCCCAACCTCCGCCTGGTACATTTCGCATCCCTGTTGCCAATTTC

Fig.52C: SEQ ID NO: 142

H3P1609(Prl-)#2.c

AATGTACCAGGCGGAGGTTGGGAGGGAATGGTGGATGGTTGGTACGGT

Fig.52D: SEQ ID NO: 143

Trình tự axit amin của PDISP/H3 Perth với vòng phân giải protein loại bớt

MAKNVAIFGLLFSLLVLVPSQIFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGEICDSP
HQILDGKNCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFNNESFNWTGVTQNG
TSSACIRRSKNSFFSRLNWLTHLNFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKDQIFLYAQASGRITVSTKRSQQTV
SPNIGSRPRVRNIPSRISIYWTIVKPGDILLINSTGNLIAPRGYFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDK
PFQNVNRITYGACPRYVKQNTLKLATGMRNVPGGGWEGMVDGWYGFRHQNSEGRGQAADLKSTQAAIDQINGKLNRLIG
KTNEKFHQIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEKTKKQLRENAEDMGNG
CFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWISFAISCFLLCVALLGFIMWACQKGNI
RCNICI*

Fig. 52E

Biểu diễn bằng sơ đồ của cấu trúc số 2039

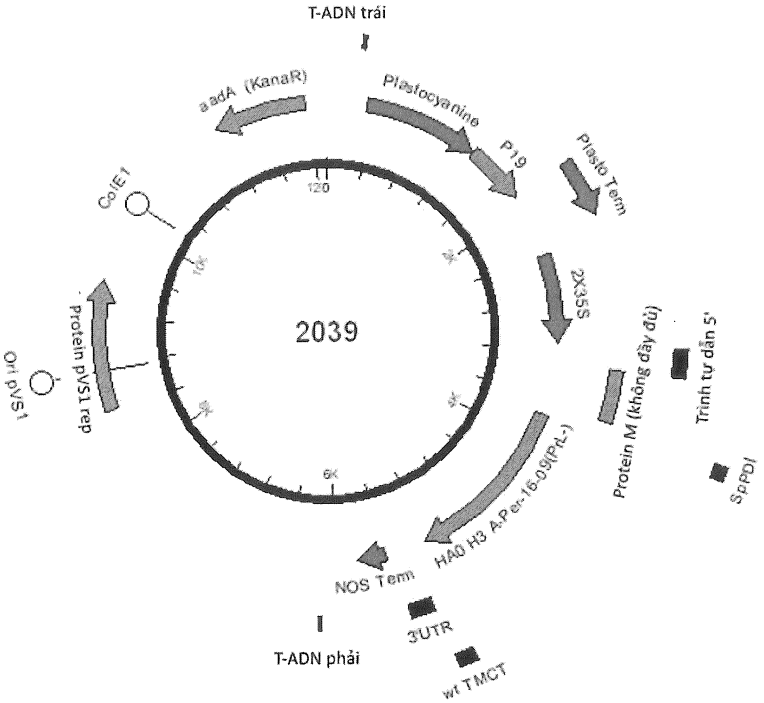
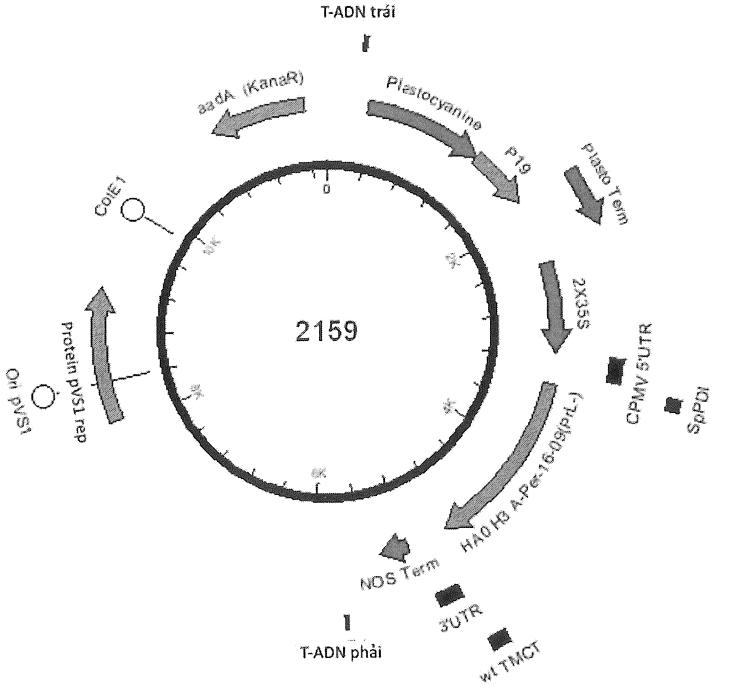


Fig. 52F

Biểu diễn bằng sơ đồ của cấu trúc số 2159



2X35S/CPMV HT+ (cấu trúc số 2230) và 160+ (cấu trúc số 2250) đối với PDISP/H3 Victoria với vòng phân giải protein loại bớt

Fig.53A: SEQ ID NO:144

Trình tự nucleotit của PDISP/H3 Victoria với vòng phân giải protein loại bớt

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGCAAAAACTTCCTGGAA
ATGACAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATAGTGAAAACAATCACGAATGACCAAA
TTGAAGTTACTAATGCTACTGAGCTGGTTCAGAATTCCTCAATAGGTGAAATATGCGACAGTCCTCATCAGATCCTTGATGGAGAA
AACTGCACACTAATAGATGCTCTATTGGGAGACCCTCAGTGTGATGGCTTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGAAG
CAAAGCCTACAGCAACTGTTACCCTTATGATGTGCCGGATTATGCCTCCCTTAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTT
TAACAATGAAAGCTTCAATTGGACTGGAGTCACTCAAAACGGAACAAGTTCTGCTTGCATAAGGAGATCTAATAATAGTTTCTTTA
GTAGATTAAATTGGTTGACCCACTTAAACTTCAAATACCCAGCATTGAACGTGACTATGCCAAACAATGAACAATTTGACAAATTG
TACATTTGGGGGGTTCACCACCCGGGTACGGACAAGGACCAAATCTTCCTGTATGCTCAATCATCAGGAAGAATCACAGTATCTAC
CAAAAGAAGCCAACAAGCTGTAATCCCGAATATCGGATCTAGACCCAGAATAAGGAATATCCCTAGCAGAATAAGCATCTATTGG
ACAATAGTAAAACCGGGAGACATACTTTTGATTAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAAATACGAAGTG
GGAAAAGCTCAATAATGAGATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCAAATGGAAGCATTCCCAATGA
CAAACCATTCCAAAATGTAAACAGGATCACATACGGGGCCTGTCCCAGATATGTTAAGCAAAGCACTCTGAAATTGGCAACAGGA
ATGCGAAATGTACCAGGCGGAGGTTGGGAGGGAATGGTGGATGGTTGGTACGGTTTCAGGCATCAAAATTCTGAGGGAAGAGG
ACAAGCAGCAGATCTCAAAAGCACTCAAGCAGCAATCGATCAAATCAATGGGAAGCTGAATCGATTGATCGGGAAAACCAACGA
GAAATTCCATCAGATTGAAAAAGAATTCTCAGAAGTCGAAGGGAGAATTCAGGACCTTGAGAAATATGTTGAGGACACTAAAATA
GATCTCTGGTCATACAACGCGGAGCTTCTTGTTGCCCTGGAGAACCAACATACAATTGATCTAACTGACTCAGAAATGAACAAACT
GTTTGAAAAAACAAAGAAGCAACTAAGGGAAAATGCTGAGGATATGGGCAATGGTTGTTTCAAAATATACCACAAATGTGACAAT
GCCTGCATAGGATCAATCAGAAATGGAACTTATGACCACGATGTATACAGAGATGAAGCATTAAACAACCGGTTCCAGATCAAGG
GAGTTGAGCTGAAGTCAGGGTACAAAGATTGGATCCTATGGATTTCCTTTGCCATATCATGTTTTTTGCTTTGTGTTGCTTTGTTGG
GGTTCATCATGTGGGCCTGCCAAAAGGGCAACATTAGGTGCAACATTTGCATTTGA

Fig.53B: SEQ ID NO: 145

H3V36111(Prl-).r

CCCTCCCAACCTCCGCCTGGTACATTTCGCATTCCTGTTGCCAATTTC

Fig.53C: SEQ ID NO: 146

H3V36111(Prl-).c

AATGTACCAGGCGGAGGTTGGGAGGGAATGGTGGATGGTTGGTACGGT

Fig.53D: SEQ ID NO: 147

Trình tự axit amin của PDISP/H3 Victoria với vòng phân giải protein loại bớt

MAKNVAIFGLLFSLLVLVPSQIFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQNSSIGEICDSPHQILDGENCTLID
ALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFNNESFNWTGVTQNGTSSACIRRSNNSFFSRLNWLTH
LNFKYPALNVTMPNNEQFDKLYIWGVHHPGTDKDQIFLYAQSSGRITVSTKRSQQAVIPNIGSRPRIRNIPSRISIYWTIVKPGDILLINST
GNLIAPRGYFKIRSGKSSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRITYGACPRYVKQSTLKLATGMRNVPGGGWEGMVDGWY
GFRHQNSEGRGQAADLKSTQAAIDQINGKLNRLIGKTNEKFHQIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDS
EMNKLFEKTKKQLRENAEDMGNGCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWISFAISCFLLCVALL
GFIMWACQKGNIRCNICI*

Fig. 53E

Biểu diễn bằng đồ thị của cấu trúc số 2230

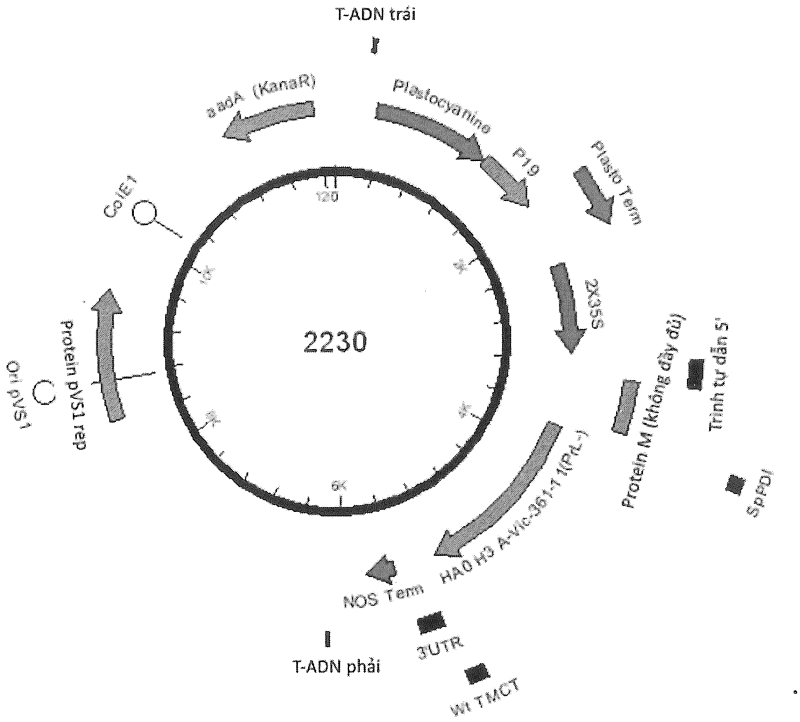
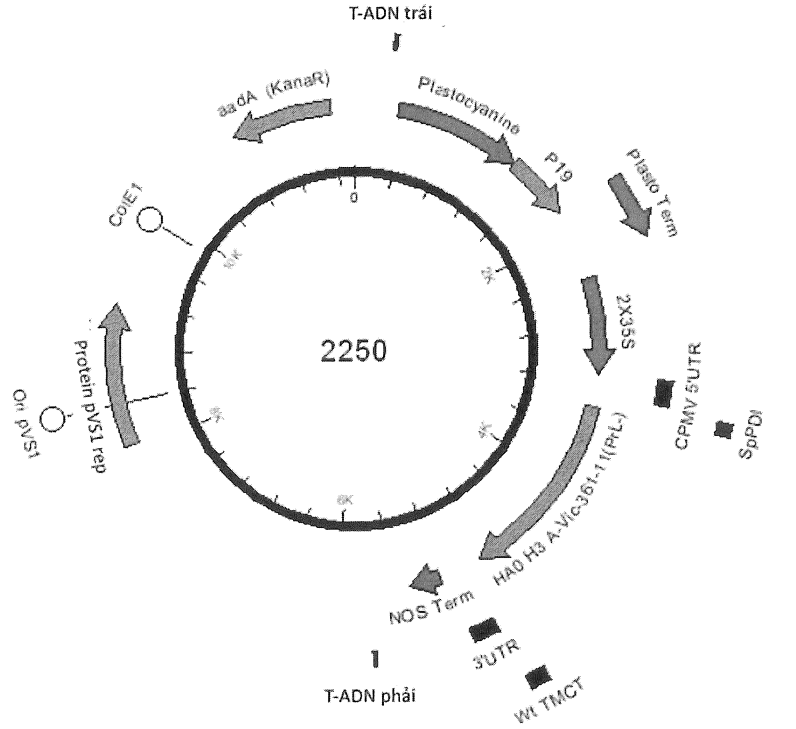


Fig. 53F

Biểu diễn bằng đồ thị của cấu trúc số 2250



2X35S/CPMV HT+/PDISP/H7 Hangzhou/NOS (cấu trúc số 2142)

Fig.54A: , SEQ ID NO: 148

Trình tự nucleotit của PDISP/H7 Hangzhou

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGACAAAATCTGCCTCG
GACATCATGCCGTGTCAAACGGAACCAAAGTAAACACATTAACTGAAAGAGGAGTGGAAGTCGTCAATGCAACTGAAACAGTGG
AACGAACAAACATCCCCAGGATCTGCTCAAAAGGGAAAAGGACAGTTGACCTCGGTCAATGTGGACTCCTGGGGACAATCACTG
GACCACCTCAATGTGACCAATTCCTAGAATTTTCAGCCGATTTAATTATTGAGAGGCGAGAAGGAAGTGATGTCTGTTATCCTGGG
AAATTCGTGAATGAAGAAGCTCTGAGGCAAATTCTCAGAGAATCAGGCGGAATTGACAAGGAAGCAATGGGATTCACATACAGT
GGAATAAGAACTAATGGAGCAACCAGTGCATGTAGGAGATCAGGATCTTCATTCTATGCAGAAATGAAATGGCTCCTGTCAAACA
CAGATAATGCTGCATTCCCGCAGATGACTAAGTCATATAAAAATACAAGAAAAAGCCCAGCTCTAATAGTATGGGGGATCCATCA
TTCCGTATCAACTGCAGAGCAAACCAAGCTATATGGGAGTGGAAACAAACTGGTGACAGTTGGGAGTTCTAATTATCAACAATCT
TTTGTACCGAGTCCAGGAGCGAGACCACAAGTTAATGGTATATCTGGAAGAATTGACTTTCATTGGCTAATGCTAAATCCCAATGA
TACAGTCACTTTCAGTTTCAATGGGGCTTTCATAGCTCCAGACCGTGCAAGCTTCCTGAGAGGAAAATCTATGGGAATCCAGAGTG
GAGTACAGGTTGATGCCAATTGTGAAGGGGACTGCTATCATAGTGGAGGGACAATAATAAGTAACTTGCCATTTCAGAACATAGA
TAGCAGGGCAGTTGGAAAATGTCCGAGATATGTTAAGCAAAGGAGTCTGCTGCTAGCAACAGGGATGAAGAATGTTCCTGAGAT
TCCAAAGGGAAGAGGCCTATTTGGTGCTATAGCGGGTTTCATTGAAAATGGATGGGAAGGCCTAATTGATGGTTGGTATGGTTTC
AGACACCAGAATGCACAGGGAGAGGGAACTGCTGCAGATTACAAAAGCACTCAATCGGCAATTGATCAAATAACAGGAAAATTA
AACCGGCTTATAGAAAAAACCAACCAACAATTTGAGTTGATCGACAATGAATTCAATGAGGTAGAGAAGCAAATCGGTAATGTGA
TAAATTGGACCAGAGATTCTATAACAGAAGTGTGGTCATACAATGCTGAACTCTTGGTAGCAATGGAGAACCAGCATACAATTGA
TCTGGCTGATTCAGAAATGGACAAACTGTACGAACGAGTGAAAAGACAGCTGAGAGAGAATGCTGAAGAAGATGGCACTGGTTG
CTTTGAAATATTTCACAAGTGTGATGATGACTGTATGGCCAGTATTAGAAATAACACCTATGATCACAGCAAATACAGGGAAGAG
GCAATGCAAAATAGAATACAGATTGACCCAGTCAAACTAAGCAGCGGCTACAAAGATGTGATACTTTGGTTTAGCTTCGGGGCAT
CATGTTTCATACTTCTAGCCATTGTAATGGGCCTTGTCTTCATATGTGTAAAGAATGGAAACATGCGGTGCACTATTTGTATATAA

Fig.54B: SEQ ID NO: 149

IF*-H7H113.s1-6r

ACTAAAGAAAATAGGCCTTTATATACAAATAGTGCACCGCATGTTTCCAT

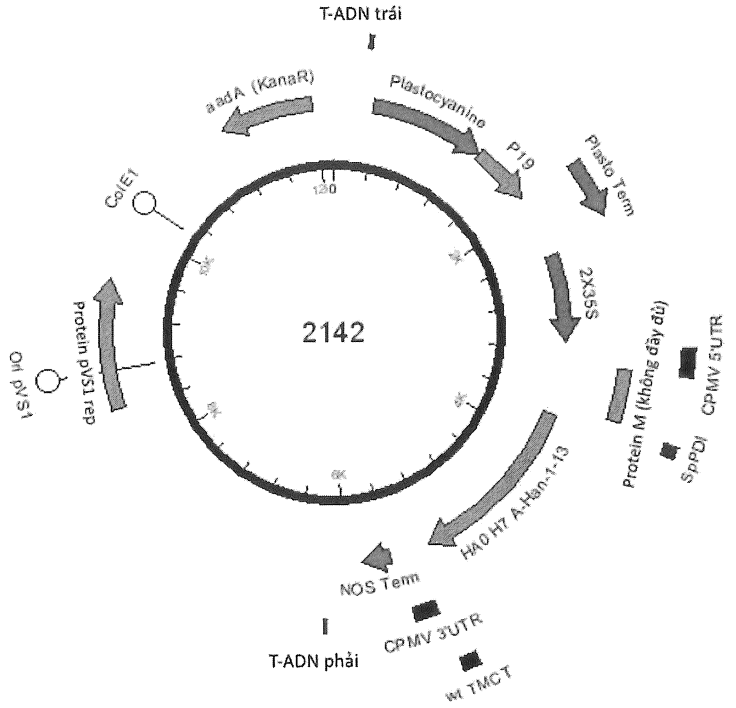
Fig.54C: SEQ ID NO: 150

Trình tự axit amin của PDISP/H7 Hangzhou

MAKNVAIFGLLFSLLVLVPSQIFADKICLGHHAVSNGTKVNTLTERGVEVVNATETVERTNIPRICSKGKRTVDLGQCGLLGTITGPPQCD
QFLEFSADLIIERREGSDVCYPGKFVNEEALRQILRESGGIDKEAMGFTYSGIRTNGATSACRRSGSSFYAEMKWLLSNTDNAAFPQMTK
SYKNTRKSPALIVWGIHHSVSTAEQTKLYGSGNKLVTVGSSNYQQSFVPSPGARPQVNGISGRIDFHWLMLNPNDTVTFSFNGAFIAPD
RASFLRGKSMGIQSGVQVDANCEGDCYHSGGTIISNLPFQNIDSRAVGKCPRYVKQRSLLLATGMKNVPEIPKGRGLFGAIAGFIENGW
EGLIDGWYGFRHQNAQGEGTAADYKSTQSAIDQITGKLNRLIEKTNQQFELIDNEFNEVEKQIGNVINWTRDSITEVWSYNAELLVAME
NQHTIDLADSEMDKLYERVKRQLRENAEEDGTGCFEIFHKCDDDCMASIRNNTYDHSKYREEAMQNRIQIDPVKLSSGYKDVILWFSFG
ASCFILLAIVMGLVFICVKNGNMRCTICI*

Fig. 54D

Biểu diễn bằng sơ đồ của cấu trúc số 2142



2X35S/CPMV HT+/PDISP/H7 Hangzhou với vòng phân giải protein loại bỏ/NOS (cấu trúc số 2152)

Fig.55A: SEQ ID NO: 151

Trình tự nucleotit của PDISP/H7 Hangzhou với vòng phân giải protein loại bỏ

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGACAAAATCTGCCTCG
GACATCATGCCGTGTCAAACGGAACCAAAGTAAACACATTAACTGAAAGAGGAGTGGAAGTCGTCAATGCAACTGAAACAGTGG
AACGAACAAACATCCCCAGGATCTGCTCAAAAGGGAAAAGGACAGTTGACCTCGGTCAATGTGGACTCCTGGGGACAATCACTG
GACCACCTCAATGTGACCAATTCCTAGAATTTTCAGCCGATTTAATTATTGAGAGGCGAGAAGGAAGTGATGTCTGTTATCCTGGG
AAATTCGTGAATGAAGAAGCTCTGAGGCAAATTCTCAGAGAATCAGGCGGAATTGACAAGGAAGCAATGGGATTCACATACAGT
GGAATAAGAACTAATGGAGCAACCAGTGCATGTAGGAGATCAGGATCTTCATTCTATGCAGAAATGAAATGGCTCCTGTCAAACA
CAGATAATGCTGCATTCCCGCAGATGACTAAGTCATATAAAAATACAAGAAAAAGCCCAGCTCTAATAGTATGGGGGATCCATCA
TTCCGTATCAACTGCAGAGCAAACCAAGCTATATGGGAGTGGAAACAAACTGGTGACAGTTGGGAGTTCTAATTATCAACAATCT
TTTGTACCGAGTCCAGGAGCGAGACCACAAGTTAATGGTATATCTGGAAGAATTGACTTTCATTGGCTAATGCTAAATCCCAATGA
TACAGTCACTTTCAGTTTCAATGGGGCTTTCATAGCTCCAGACCGTGCAAGCTTCCTGAGAGGAAAATCTATGGGAATCCAGAGTG
GAGTACAGGTTGATGCCAATTGTGAAGGGGACTGCTATCATAGTGGAGGGACAATAATAAGTAACTTGCCATTTCAGAACATAGA
TAGCAGGGCAGTTGGAAAATGTCCGAGATATGTTAAGCAAAGGAGTCTGCTGCTAGCAACAGGGATGAAGAATGTTCCTGGCGG
ATGGGAAGGCCTAATTGATGGTTGGTATGGTTTCAGACACCAGAATGCACAGGGAGAGGGAACTGCTGCAGATTACAAAAGCAC
TCAATCGGCAATTGATCAAATAACAGGAAAATTAAACCGGCTTATAGAAAAAACCAACCAACAATTTGAGTTGATCGACAATGAA
TTCAATGAGGTAGAGAAGCAAATCGGTAATGTGATAAATTGGACCAGAGATTCTATAACAGAAGTGTGGTCATACAATGCTGAAC
TCTTGGTAGCAATGGAGAACCAGCATACAATTGATCTGGCTGATTCAGAAATGGACAAACTGTACGAACGAGTGAAAAGACAGCT
GAGAGAGAATGCTGAAGAAGATGGCACTGGTTGCTTTGAAATATTTCACAAGTGTGATGATGACTGTATGGCCAGTATTAGAAAT
AACACCTATGATCACAGCAAATACAGGGAAGAGGCAATGCAAAATAGAATACAGATTGACCCAGTCAAACTAAGCAGCGGCTAC
AAAGATGTGATACTTTGGTTTAGCTTCGGGGCATCATGTTTCATACTTCTAGCCATTGTAATGGGCCTTGTCTTCATATGTGTAAAG
AATGGAAACATGCGGTGCACTATTTGTATATAA

Fig.55B: SEQ ID NO: 152

H7H113(PrL-).r

CCTTCCCATCCGCCAGGAACATTCTTCATCCCTGTTGCTAGCAGCAGAC

Fig.55C: SEQ ID NO: 153

H7H113(PrL-).c

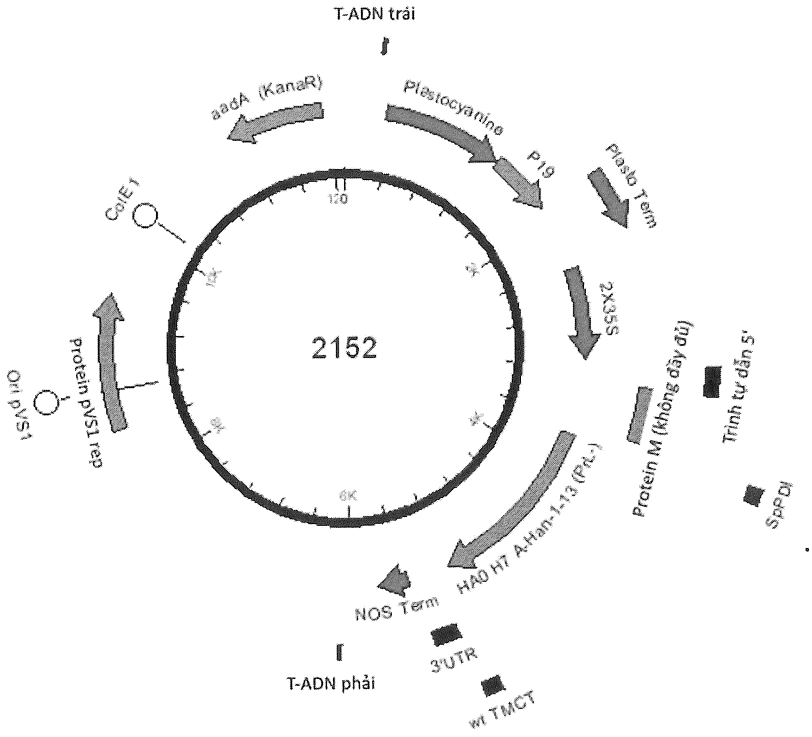
AGAATGTTCCTGGCGGATGGGAAGGCCTAATTGATGGTTGGTATGGTT

Fig.55D: SEQ ID NO: 154

Trình tự axit amin của PDISP/H7 Hangzhou với vòng phân giải protein loại bót

MAKNVAIFGLLFSLLVLVPSQIFADKICLGHHAVSNGTKVNTLTERGVEVVNATETVERTNIPRICSKGKRTVDLGQCGLLGTITGPPQCD
QFLEFSADLIIERREGSDVCYPGKFVNEEALRQILRESGGIDKEAMGFTYSGIRTNGATSACRRSGSSFYAEMKWLLSNTDNAAFPQMTK
SYKNTRKSPALIVWGIHHSVSTAEQTKLYGSGNKLVTVGSSNYQQSFVPSPGARPQVNGISGRIDFHWLMLNPNDTVTFSFNGAFIAPD
RASFLRGKSMGIQSGVQVDANCEGDCYHSGGTIISNLPFQNIDSRAVGKCPRYVKQRSLLLATGMKNVPGGWEGLIDGWYGFRHQNA
QGEGTAADYKSTQSAIDQITGKLNRLIEKTNQQFELIDNEFNEVEKQIGNVINWTRDSITEVWSYNAELLVAMENQHTIDLADSEMDKL
YERVKRQLRENAEEDGTGCFEIFHKCDDDCMASIRNNTYDHSKYREEAMQNRIQIDPVKLSSGYKDVILWFSFGASCFILLAIVMGLVFIC
VKNGNMRCTICI*

Fig. 55E
Biểu diễn bằng sơ đồ của cấu trúc số 2152



2X35S/CPMV HT+ (cấu trúc số 2224) và 160+ (cấu trúc số 2226) đối với PDISP/H9 Hong Kong

Fig.56A: SEQ ID NO: 155

Trình tự nucleotit của PDISP/H9 Hong Kong

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGATAAAATCTGCATCG
GCCACCAGTCAACAAACTCCACAGAAACTGTGGACACGCTAACAGAAACCAATGTTCCTGTGACACATGCCAAAGAATTGCTCCAC
ACAGAGCATAATGGAATGCTGTGTGCAACAAGCCTGGGACATCCCCTCATTCTAGACACATGCACTATTGAAGGACTAGTCTATG
GCAACCCTTCTTGTGACCTGCTGTTGGGAGGAAGAGAATGGTCCTACATCGTCGAAAGATCATCAGCTGTAAATGGAACGTGTTA
CCCTGGGAATGTAGAAAACCTAGAGGAACTCAGGACACTTTTTAGTTCCGCTAGTTCCTACCAAAGAATCCAAATCTTCCCAGACA
CAACCTGGAATGTGACTTACACTGGAACAAGCAGAGCATGTTCAGGTTCATTCTACAGGAGTATGAGATGGCTGACTCAAAAGAG
CGGTTTTTACCCTGTTCAAGACGCCCAATACACAAATAACAGGGGAAAGAGCATTCTTTTCGTGTGGGGCATACATCACCCACCCA
CCTATACCGAGCAAACAAATTTGTACATAAGAAACGACACAACAACAAGCGTGACAACAGAAGATTTGAATAGGACCTTCAAACC
AGTGATAGGGCCAAGGCCCCTTGTCAATGGTCTGCAGGGAAGAATTGATTATTATTGGTCGGTACTAAAACCAGGCCAAACATTG
CGAGTACGATCCAATGGGAATCTAATTGCTCCATGGTATGGACACGTTCTTTCAGGAGGGAGCCATGGAAGAATCCTGAAGACTG
ATTTAAAAGGTGGTAATTGTGTAGTGCAATGTCAGACTGAAAAAGGTGGCTTAAACAGTACATTGCCATTCCACAATATCAGTAAA
TATGCATTTGGAACCTGCCCCAAATATGTAAGAGTTAATAGTCTCAAACTGGCAGTCGGTCTGAGGAACGTGCCTGCTAGATCAA
GTAGAGGACTATTTGGAGCCATAGCTGGATTCATAGAAGGAGGTTGGCCAGGACTAGTCGCTGGCTGGTATGGTTTCCAGCATTC
AAATGATCAAGGGGTTGGTATGGCTGCAGATAGGGATTCAACTCAAAAGGCAATTGATAAAATAACATCCAAGGTGAATAATATA
GTCGACAAGATGAACAAGCAATATGAAATAATTGATCATGAATTTAGTGAGGTTGAAACTAGACTCAATATGATCAATAATAAGA
TTGATGACCAAATACAAGACGTATGGGCATATAATGCAGAATTGCTAGTACTACTTGAAAATCAAAAAACACTCGATGAGCATGA
TGCGAACGTGAACAATCTATATAACAAGGTGAAGAGGGCACTGGGCTCCAATGCTATGGAAGATGGGAAAGGCTGTTTCGAGCT
ATACCATAAATGTGATGATCAGTGCATGGAAACAATTCGGAACGGGACCTATAATAGGAGAAAGTATAGAGAGGAATCAAGACT
AGAAAGGCAGAAAATAGAGGGGGTTAAGCTGGAATCTGAGGGAACTTACAAAATCCTCACCATTTATTCGACTGTCGCCTCATCT
CTTGTGCTTGCAATGGGGTTTGCTGCCTTCCTGTTCTGGGCCATGTCCAATGGATCTTGCAGATGCAACATTTGTATATAA

Fig.56B: SEQ ID NO: 156

IF**-H9HK107399.S1-6r

ACTAAAGAAAATAGGCCTTTATATACAAATGTTGCATCTGCAAGATCCAT

Fig.56C: SEQ ID NO: 157

Trình tự axit amin của PDISP/H9 Hong Kong

MAKNVAIFGLLFSLLVLVPSQIFADKICIGHQSTNSTETVDTLTETNVPVTHAKELLHTEHNGMLCATSLGHPLILDTCTIEGLVYGNPSCD
LLLGGREWSYIVERSSAVNGTCYPGNVENLEELRTLFSSASSYQRIQIFPDTTWNVTYTGTSRACSGSFYRSMRWLTQKSGFYPVQDAQ
YTNNRGKSILFVWGIHHPPTYTEQTNLYIRNDTTTSVTTEDLNRTFKPVIGPRPLVNGLQGRIDYYWSVLKPGQTLRVRSNGNLIAPWYG
HVLSGGSHGRILKTDLKGGNCVVQCQTEKGGLNSTLPFHNISKYAFGTCPKYVRVNSLKLAVGLRNVPARSSRGLFGAIAGFIEGGWPGL
VAGWYGFQHSNDQGVGMAADRDSTQKAIDKITSKVNNIVDKMNKQYEIIDHEFSEVETRLNMINNKIDDQIQDVWAYNAELLVLLEN
QKTLDEHDANVNNLYNKVKRALGSNAMEDGKGCFELYHKCDDQCMETIRNGTYNRRKYREESRLERQKIEGVKLESEGTYKILTIYSTVA
SSLVLAMGFAAFLFWAMSNGSCRCNICI*

Fig. 56D

Biểu diễn bằng sơ đồ của cấu trúc số 2224

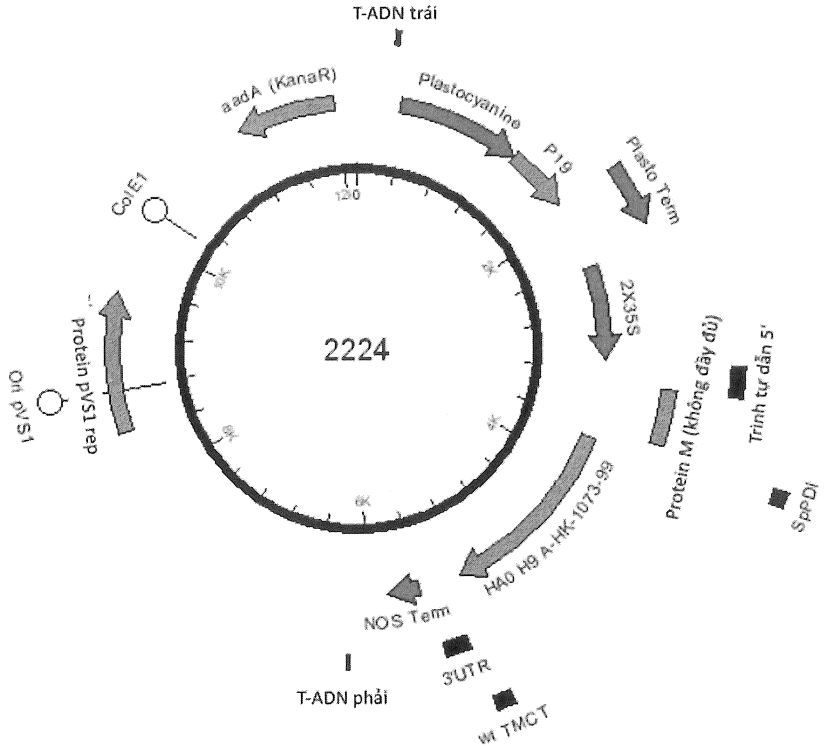
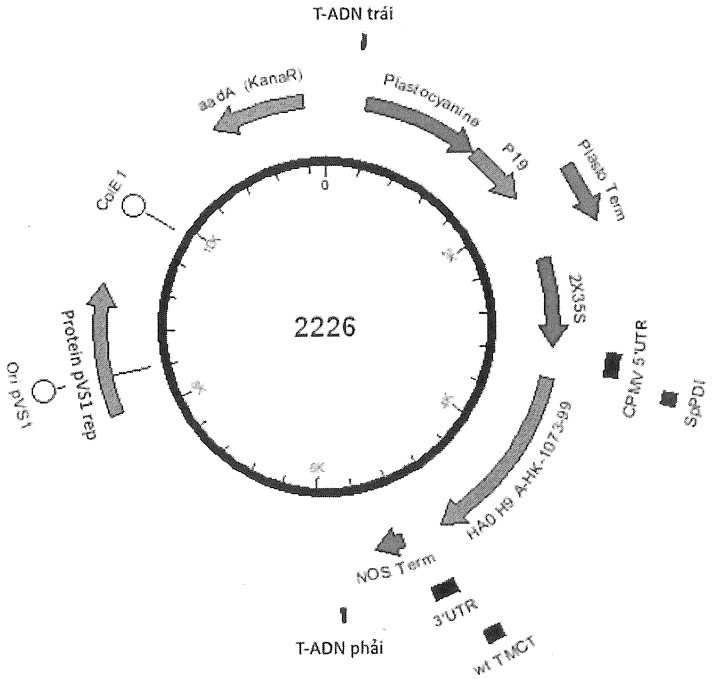


Fig. 56E

Biểu diễn bằng sơ đồ của cấu trúc 2226



2X35S/CPMV HT+ (cấu trúc số 2225) và 160+ (cấu trúc số 2227) đối với PDISP/H9 Hong Kong với vòng phân giải protein loại bớt

Fig.57A: SEQ ID NO: 158

Trình tự nucleotit của PDISP/H9 Hong Kong với vòng phân giải protein loại bớt

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGATAAAATCTGCATCG
GCCACCAGTCAACAAACTCCACAGAAACTGTGGACACGCTAACAGAAACCAATGTTCCTGTGACACATGCCAAAGAATTGCTCCAC
ACAGAGCATAATGGAATGCTGTGTGCAACAAGCCTGGGACATCCCCTCATTCTAGACACATGCACTATTGAAGGACTAGTCTATG
GCAACCCTTCTTGTGACCTGCTGTTGGGAGGAAGAGAATGGTCCTACATCGTCGAAAGATCATCAGCTGTAAATGGAACGTGTTA
CCCTGGGAATGTAGAAAACCTAGAGGAACTCAGGACACTTTTTAGTTCCGCTAGTTCCTACCAAAGAATCCAAATCTTCCCAGACA
CAACCTGGAATGTGACTTACACTGGAACAAGCAGAGCATGTTCAGGTTCATTCTACAGGAGTATGAGATGGCTGACTCAAAAGAG
CGGTTTTTACCCTGTTCAAGACGCCCAATACACAAATAACAGGGGAAAGAGCATTCTTTTCGTGTGGGGCATACATCACCCACCCA
CCTATACCGAGCAAACAAATTTGTACATAAGAAACGACACAACAACAAGCGTGACAACAGAAGATTTGAATAGGACCTTCAAACC
AGTGATAGGGCCAAGGCCCCTTGTCAATGGTCTGCAGGGAAGAATTGATTATTATTGGTCGGTACTAAAACCAGGCCAAACATTG
CGAGTACGATCCAATGGGAATCTAATTGCTCCATGGTATGGACACGTTCTTTCAGGAGGGAGCCATGGAAGAATCCTGAAGACTG
ATTTAAAAGGTGGTAATTGTGTAGTGCAATGTCAGACTGAAAAAGGTGGCTTAAACAGTACATTGCCATTCCACAATATCAGTAAA
TATGCATTTGGAACCTGCCCCAAATATGTAAGAGTTAATAGTCTCAAACTGGCAGTCGGTCTGAGGAACGTGCCTGGCGGAGGTT
GGCCAGGACTAGTCGCTGGCTGGTATGGTTTCCAGCATTCAAATGATCAAGGGGTTGGTATGGCTGCAGATAGGGATTCAACTCA
AAAGGCAATTGATAAAATAACATCCAAGGTGAATAATATAGTCGACAAGATGAACAAGCAATATGAAATAATTGATCATGAATTT
AGTGAGGTTGAAACTAGACTCAATATGATCAATAATAAGATTGATGACCAAATACAAGACGTATGGGCATATAATGCAGAATTGC
TAGTACTACTTGAAAATCAAAAAACACTCGATGAGCATGATGCGAACGTGAACAATCTATATAACAAGGTGAAGAGGGCACTGG
GCTCCAATGCTATGGAAGATGGGAAAGGCTGTTTCGAGCTATACCATAAATGTGATGATCAGTGCATGGAAACAATTCGGAACGG
GACCTATAATAGGAGAAAGTATAGAGAGGAATCAAGACTAGAAAGGCAGAAAATAGAGGGGGTTAAGCTGGAATCTGAGGGAA
CTTACAAAATCCTCACCATTTATTCGACTGTCGCCTCATCTCTTGTGCTTGCAATGGGGTTTGCTGCCTTCCTGTTCTGGGCCATGTC
CAATGGATCTTGCAGATGCAACATTTGTATATAA

Fig.57B: SEQ ID NO: 159

H9HK107399(Prl-).r

GTCCTGGCCAACCTCCGCCAGGCACGTTCCTCAGACCGACTGCCAGTT

Fig.57C: SEQ ID NO: 160

H9HK107399(Prl-).c

GGAACGTGCCTGGCGGAGGTTGGCCAGGACTAGTCGCTGGCTGGTATG

Fig.57D: SEQ ID NO: 161

Trình tự axit amin của PDISP/H9 Hong Kong với vòng phân giải protein loại bớt

MAKNVAIFGLLFSLLVLVPSQIFADKICIGHQSTNSTETVDTLTETNVPVTHAKELLHTEHNGMLCATSLGHPLILDTCTIEGLVYGNPSCD
LLLGGREWSYIVERSSAVNGTCYPGNVENLEELRTLFSSASSYQRIQIFPDTTWNVTYTGTSRACSGSFYRSMRWLTQKSGFYPVQDAQ
YTNNRGKSILFVWGIHHPPTYTEQTNLYIRNDTTTSVTTEDLNRTFKPVIGPRPLVNGLQGRIDYYWSVLKPGQTLRVRSNGNLIAPWYG
HVLSGGSHGRILKTDLKGGNCVVQCQTEKGGLNSTLPFHNISKYAFGTCPKYVRVNSLKLAVGLRNVPGGGWPGLVAGWYGFQHSND
QGVGMAADRDSTQKAIDKITSKVNNIVDKMNKQYEIIDHEFSEVETRLNMINNKIDDQIQDVWAYNAELLVLLENQKTLDEHDANVN
NLYNKVKRALGSNAMEDGKGCFELYHKCDDQCMETIRNGTYNRRKYREESRLERQKIEGVKLESEGTYKILTIYSTVASSLVLAMGFAAF
LFWAMSNGSCRCNICI*

Fig. 57E
Biểu diễn bằng đồ thị của cấu trúc số 2225

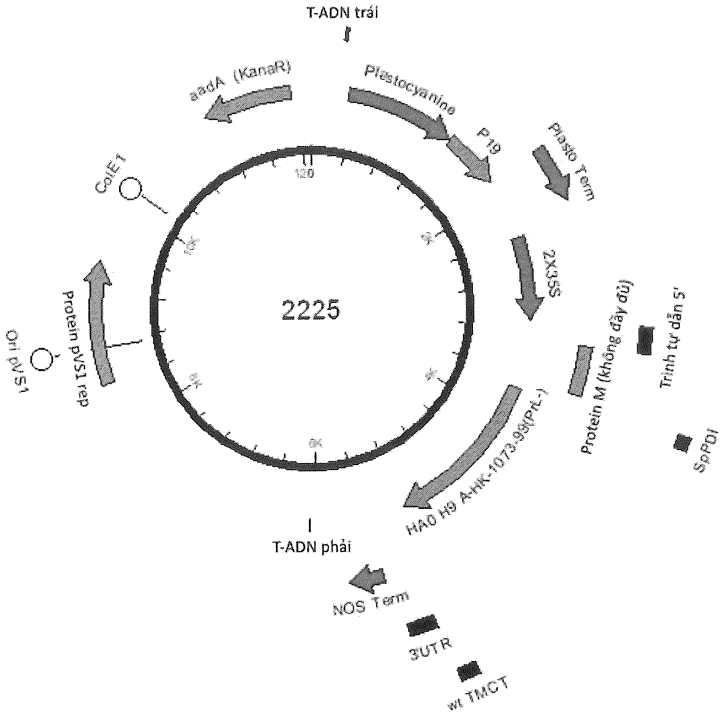
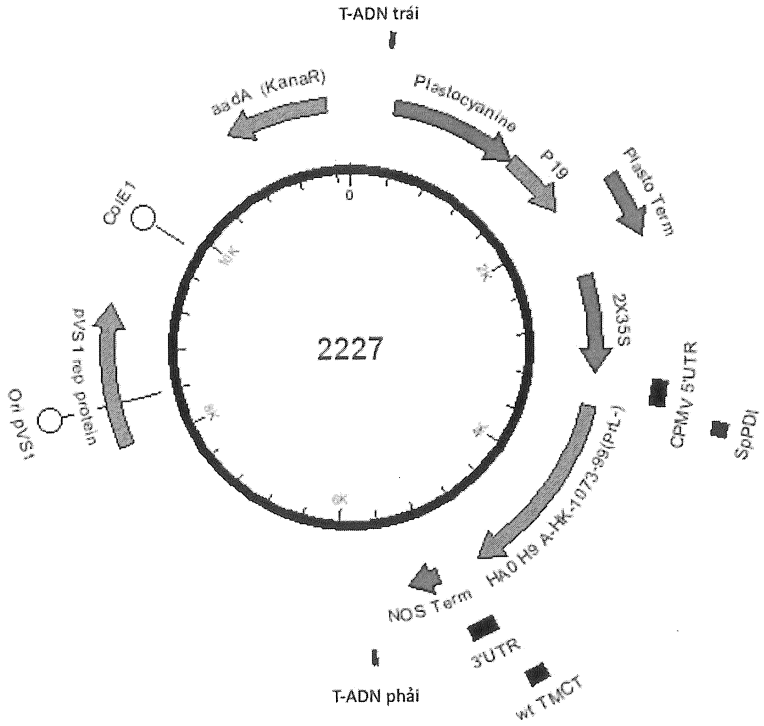


Fig. 57F
Biểu diễn bằng sơ đồ của cấu trúc số 2227



2X35S/CPMV 160+/PDISP/HA B Malaysia/NOS (cấu trúc số 2013)

Fig.58A: SEQ ID NO: 162

Trình tự nucleotit của PDISP/HA B Malaysia

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGATCGAATCTGCACTG
GGATAACATCGTCAAACTCACCACATGTTGTCAAAACTGCTACTCAAGGGGAGGTCAATGTGACTGGTGTAATACCACTGACAAC
AACACCCACCAAATCTCATTTTGCAAATCTCAAAGGAACAGAAACCAGAGGGAAACTATGCCCAAAATGCCTCAACTGCACAGATC
TGGACGTGGCCTTGGGCAGACCAAAATGCACGGGGAACATACCCTCGGCAAGAGTTTCAATACTCCATGAAGTCAGACCTGTTAC
ATCTGGGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCCTAAACTTCTCAGAGGATACGAACATATCAGGTTAT
CAACTCATAACGTTATCAATGCAGAAAATGCACCAGGAGGACCCTACAAAATTGGAACCTCAGGGTCTTGCCCTAACGTTACCAAT
GGAAACGGATTTTTCGCAACAATGGCTTGGGCCGTCCCAAAAAACGACAACAACAAAACAGCAACAAATTCATTAACAATAGAAG
TACCATACATTTGTACAGAAGGAGAAGACCAAATTACCGTTTGGGGGTTCCACTCTGATAACGAAACCCAAATGGCAAAGCTCTAT
GGGGACTCAAAGCCCCAGAAGTTCACCTCATCTGCCAACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCAAATCA
AACAGAAGACGGAGGACTACCACAAAGCGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGGGAAAACAGGAACAATTAC
CTATCAAAGAGGTATTTTATTGCCTCAAAAAGTGTGGTGCGCAAGTGGCAGGAGCAAGGTAATAAAAGGATCGTTGCCTTTAATT
GGAGAAGCAGATTGCCTCCACGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATA
GGAAATTGCCCAATATGGGTGAAAACACCCTTGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGCAAAACTATTAAAGGAAA
GGGGTTTCTTCGGAGCTATTGCTGGTTTCTTAGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCCCATG
GGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCACTCAAGAGGCCATAAACAAGATAACAAAAAATCTCAACTCTTTGA
GTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCGGTGCCATGGATGAACTCCACAACGAAATACTAGAACTAGACGAGAAAG
TGGATGATCTCAGAGCTGATACAATAAGCTCACAAATAGAACTCGCAGTCCTGCTTTCCAATGAAGGAATAATAAACAGTGAAGA
TGAGCATCTCTTGGCGCTTGAAAGAAAGCTGAAGAAAATGCTGGGCCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTGAAACC
AAACACAAGTGCAACCAGACCTGTCTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCA
CTGAATATTACTGCTGCATCTTTAAATGACGATGGATTGGATAATCATACTATACTGCTTTACTACTCAACTGCTGCCTCCAGTTTG
GCTGTAACATTGATGATAGCTATCTTTGTTGTTTATATGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGTCTATAA

Fig.58B: SEQ ID NO: 163

IF**-HBM250604.S1-6r

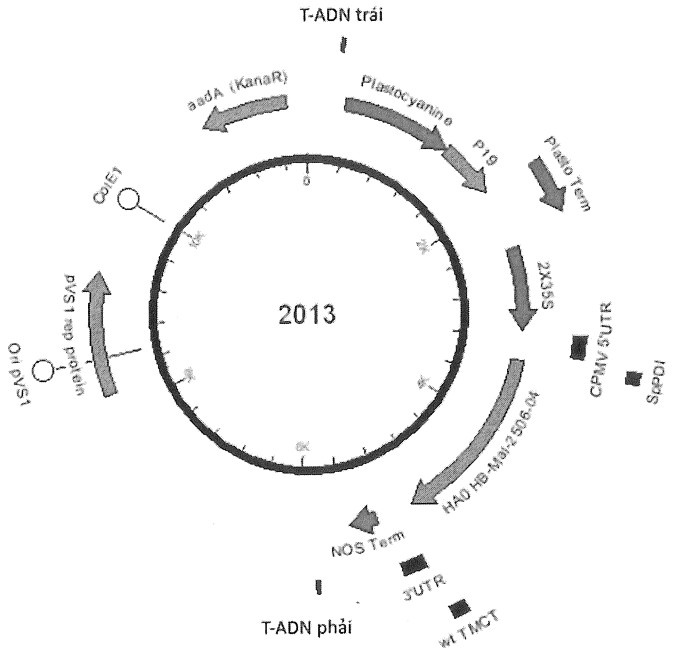
ACTAAAGAAAATAGGCCTTTATAGACAGATGGAGCAAGAAACATTGTCTC

Fig.58C: SEQ ID NO: 164

Trình tự axit amin của PDISP/HA B Malaysia

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKGTETRGKLCPKCLNCTDLDVALG
RPKCTGNIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPKLLRGYEHIRLSTHNVINAENAPGGPYKIGTSGSCPNVTNGNGFFATMAW
AVPKNDNNKTATNSLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQTEDGGLPQSGRI
VVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTK
YRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEIL
ELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSLN
ITAASLNDDGLDNHTILLYYSTAASSLAVTLMIAIFVVYMVSRDNVSCSICL*

Fig. 58D
Biểu diễn bằng sơ đồ của cấu trúc số 2013



2X35S/CPMV 160+/PDISP/HA B Malaysia với vòng phân giải protein loại bớt/NOS (cấu trúc số 2014)

Fig.59A: SEQ ID NO:165

Trình tự nucleotit của PDISP/HA B Malaysia với vòng phân giải protein

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGATCGAA
TCTGCACTGGGATAACATCGTCAAACTCACCACATGTTGTCAAAACTGCTACTCAAGGGGAGGTCAATGTGACTGGTGT
AATACCACTGACAACAACACCCACCAAATCTCATTTTGCAAATCTCAAAGGAACAGAAACCAGAGGGAAACTATGCCCA
AAATGCCTCAACTGCACAGATCTGGACGTGGCCTTGGGCAGACCAAAATGCACGGGGAACATACCCTCGGCAAGAGTTT
CAATACTCCATGAAGTCAGACCTGTTACATCTGGGTGCTTTCCTATAATGCACGACAGAACAAAAATTAGACAGCTGCC
TAAACTTCTCAGAGGATACGAACATATCAGGTTATCAACTCATAACGTTATCAATGCAGAAAATGCACCAGGAGGACCC
TACAAAATTGGAACCTCAGGGTCTTGCCCTAACGTTACCAATGGAAACGGATTTTTCGCAACAATGGCTTGGGCCGTCC
CAAAAAACGACAACAACAAAACAGCAACAAATTCATTAACAATAGAAGTACCATACATTTGTACAGAAGGAGAAGACCA
AATTACCGTTTGGGGGTTCCACTCTGATAACGAAACCCAAATGGCAAAGCTCTATGGGGACTCAAAGCCCCAGAAGTTC
ACCTCATCTGCCAACGGAGTGACCACACATTACGTTTCACAGATTGGTGGCTTCCCAAATCAAACAGAAGACGGAGGAC
TACCACAAAGCGGTAGAATTGTTGTTGATTACATGGTGCAAAAATCTGGGAAAACAGGAACAATTACCTATCAAAGAGG
TATTTTATTGCCTCAAAAAGTGTGGTGCGCAAGTGGCAGGAGCAAGGTAATAAAAGGATCGTTGCCTTTAATTGGAGAA
GCAGATTGCCTCCACGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGGGAACATGCAAAGGCCATAG
GAAATTGCCCAATATGGGTGAAAACACCCTTGAAGCTGGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGCGA
AGGAATGATTGCAGGTTGGCACGGATACACATCCCATGGGGCACATGGAGTAGCGGTGGCAGCAGACCTTAAGAGCACT
CAAGAGGCCATAAACAAGATAACAAAAAATCTCAACTCTTTGAGTGAGCTGGAAGTAAAGAATCTTCAAAGACTAAGCG
GTGCCATGGATGAACTCCACAACGAAATACTAGAACTAGACGAGAAAGTGGATGATCTCAGAGCTGATACAATAAGCTC
ACAAATAGAACTCGCAGTCCTGCTTTCCAATGAAGGAATAATAAACAGTGAAGATGAGCATCTCTTGGCGCTTGAAAGA
AAGCTGAAGAAAATGCTGGGCCCCTCTGCTGTAGAGATAGGGAATGGATGCTTTGAAACCAAACACAAGTGCAACCAGA
CCTGTCTCGACAGAATAGCTGCTGGTACCTTTGATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCACTGAATATTAC
TGCTGCATCTTTAAATGACGATGGATTGGATAATCATACTATACTGCTTTACTACTCAACTGCTGCCTCCAGTTTGGCT
GTAACATTGATGATAGCTATCTTTGTTGTTTATATGGTCTCCAGAGACAATGTTTCTTGCTCCATCTGTCTATAA

Fig.59B: SEQ ID NO: 166

HBM250604(PrL-).r

CATTCCTTCCCATCCTCCACCAGGAGGTCTATATTTGGTTCCATTGGC

Fig.59C: SEQ ID NO: 167

HBM250604(PrL-).c

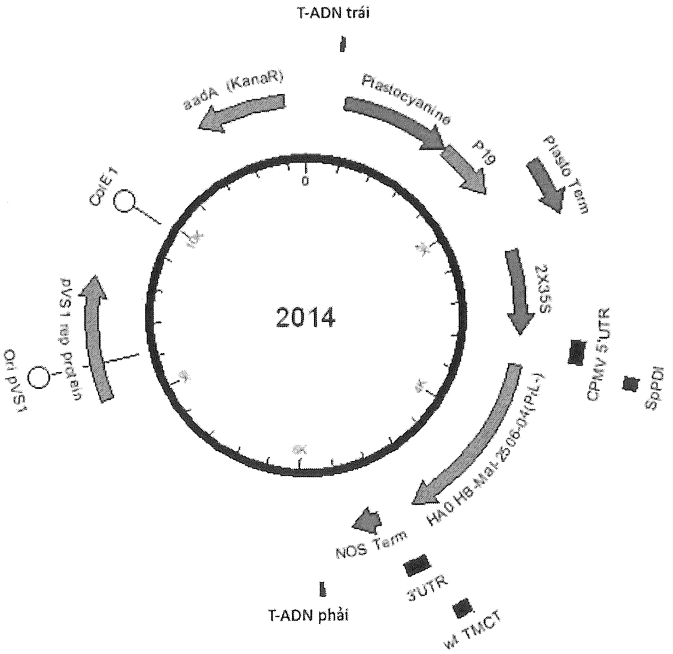
AGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGA

Fig.59D: SEQ ID NO: 168

Trình tự axit amin của PDISP/HA B Malaysia với vòng phân giải protein loại bớt

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLKGTETRGKLCPKCLNCTDLDVALG
RPKCTGNIPSARVSILHEVRPVTSGCFPIMHDRTKIRQLPKLLRGYEHIRLSTHNVINAENAPGGPYKIGTSGSCPNVTNGNGFFATMAW
AVPKNDNNKTATNSLTIEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQTEDGGLPQSGRI
VVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTK
YRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISS
QIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSLNITAASLNDDGLDNHT
ILLYYSTAASSLAVTLMIAIFVVYMVSRDNVSCSICL*

Fig. 59E
Biểu diễn bằng sơ đồ của cấu trúc 2014



2X35S/CPMV HT (cấu trúc số 2070), HT+ (cấu trúc số 2080) và 160+ (-Mprot) (cấu trúc số 2090) đối với PDISP/HA B Massachusetts

Fig.60A: SEQ ID NO:169

Trình tự Nucleotit của PDISP/HA B Massachusetts

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGATCGAATCTGCACTG
GGATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCTACTCAAGGGGAGGTCAATGTGACTGGTGTGATACCACTAACAAC
AACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAAGACCAGAGGGAAACTATGCCCAGACTGTCTCAACTGTACAGATC
TGGATGTGGCCCTGGGCAGGCCAATGTGTGTGGGAACTACACCTTCTGCGAAAGCTTCAATACTTCACGAAGTCAGACCTGTTAC
ATCCGGGTGCTTCCCTATAATGCACGACAGAACAAAAATCAGGCAACTAGCCAATCTTCTCAGAGGATATGAAAATATCAGGTTAT

CAACCCAAAACGTTATCGATGCAGAAAAGGCACCAGGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAG
TAAAAGCGGATTTTTCGCAACAATGGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCATTAACAGTAGAAGTA
CCATACATTTGTGCAGAAGGGGAAGACCAAATTACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATG
GAGACTCAAATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACATTATGTTTCTCAGATTGGCGGCTTCCCAGATCAAA
CAGAAGACGGAGGACTACCACAAAGCGGCAGAATTGTCGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCT
ATCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCTTTAATTGG
TGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGG
AAATTGCCCAATATGGGTGAAAACACCTTTGAAGCTTGCCAATGGAACCAAATATAGACCTCCTGCAAAACTATTAAAGGAAAGG
GGTTTCTTCGGAGCTATTGCTGGTTTCCTAGAAGGAGGATGGGAAGGAATGATTGCAGGTTGGCACGGATACACATCTCACGGA
GCACATGGAGTGGCAGTTGCTGCAGACCTTAAGAGCACACAAGAAGCTATAAACAAGATAACAAAAAATCTCAACTCTTTGAGTG
AGCTAGAAGTAAAGAATCTTCAAAGGCTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGG
ATGACCTCAGAGCTGACACTATAAGTTCACAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGACGA
GCATCTATTGGCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGACATAGGAAATGGATGCTTCGAAACCAAA
CACAAATGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAGTTTTCTCTCCCCACTTTTGATTCATTG
AACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTTCTAGTTTGGCTG
TAACATTGATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCCATCTGTCTATAA

Fig.60B: SEQ ID NO:170

Trình tự axit amin của PDISP/HA B Massachusetts

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTKTRGKLCPDCLNCTDLDVALG
RPMCVGTTPSAKASILHEVRPVTSGCFPIMHDRTKIRQLANLLRGYENIRLSTQNVIDAEKAPGGPYRLGTSGSCPNATSKSGFFATMA
WAVPKDNNKNATNPLTVEVPYICAEGEDQITVWGFHSDNKTQMKNLYGDSNPQKFTSSANGVTTHYVSQIGGFPDQTEDGGLPQSG
RIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEADCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLAN
GTKYRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELH
NEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTF
DSLNITAASLNDDGLDNHTILLYYSTAASSLAVTLMLAIFIVYMVSRDNVSCSICL*

Fig. 60C
Biểu diễn bằng sơ đồ của cấu trúc số 2070

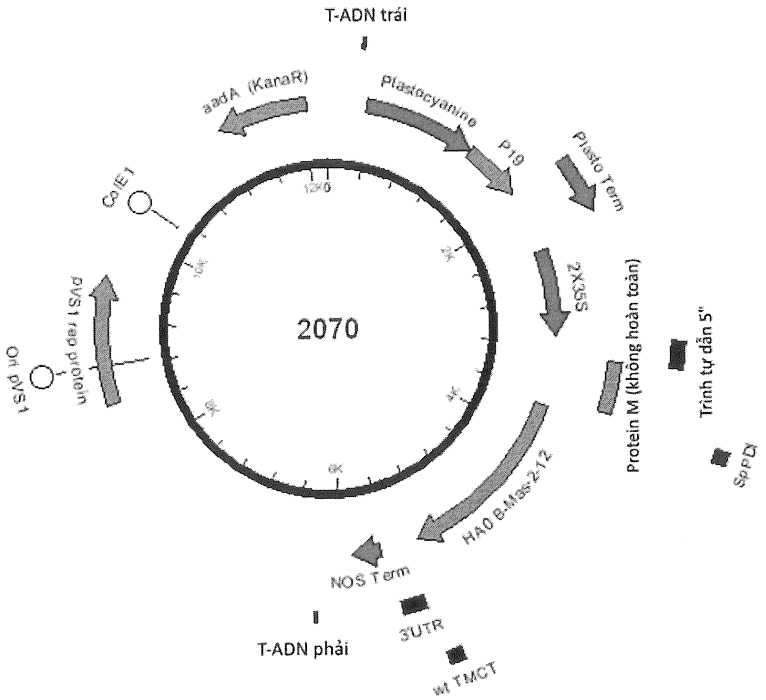


Fig. 60D
Biểu diễn bằng sơ đồ của cấu trúc số 2080

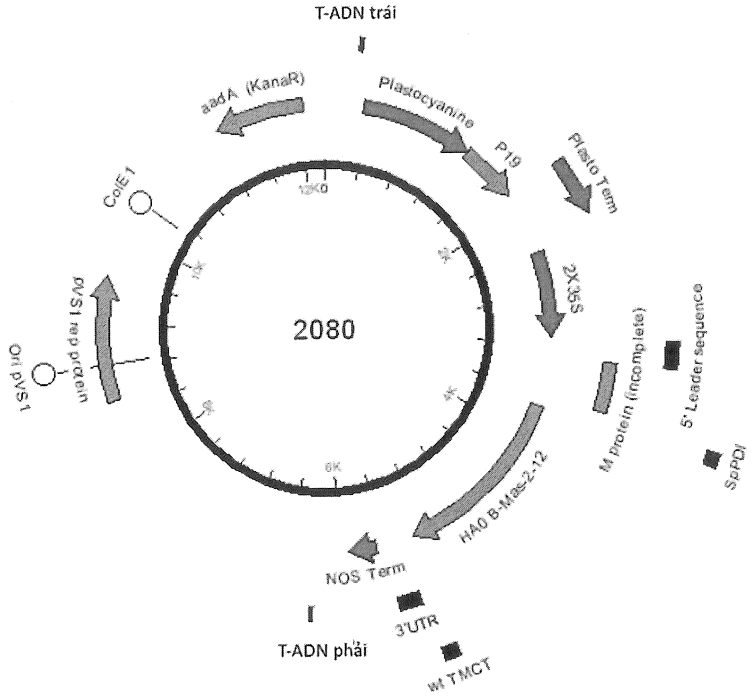
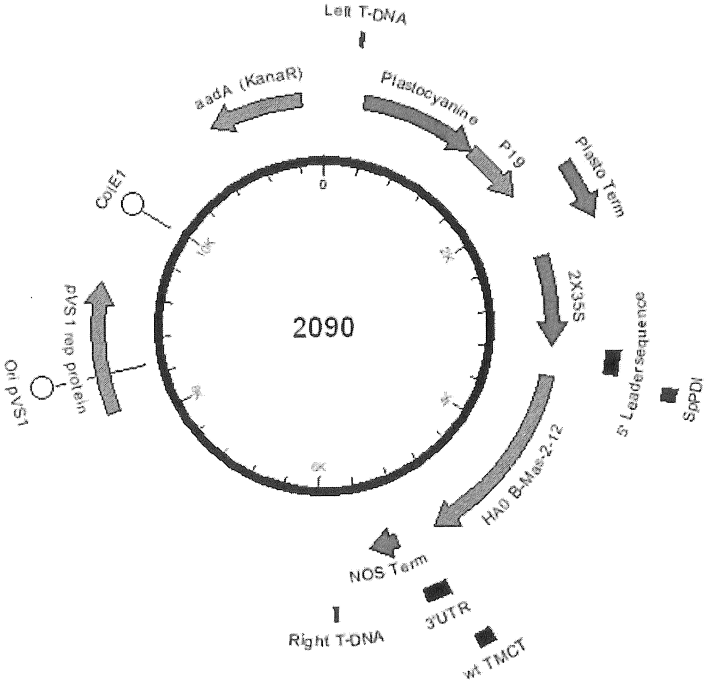


Fig. 60E
Biểu diễn bằng sơ đồ của cấu trúc số 2090



2X35S/CPMV HT+/PDISP/B Floria ới vòng phân giải protein loại bớt/NOS (cấu trúc số 2102) và 2X35S/CPMV HT+/BeYDV/PDISP/B Florida với vòng phân giải protein loại bớt/NOS (cấu trúc số 2014)

Fig 61A SEQ ID NO: 190

HBF406(PrL-).r (cấu trúc 2102 và 2104)

TCCTTCCCATCCTCCACCAGGAGGTCTATATTTGGTTCCATTGGCGAGCTTCAAAG

Fig 61B SEQ ID NO: 191

HBF406(PrL-).c (cấu trúc 2102 và 2104)

ATATAGACCTCCTGGTGGAGGATGGGAAGGAATGATTGCAGGCTGGCACGGA

Fig 61C SEQ ID NO: 192

IF*-HBF406.s1-6r (cấu trúc 2102 và 2104)

ACTAAAGAAAATAGGCCTTTATAGACAGATGGAGCATGAAACGTTGTCTC

Fig 61D SEQ ID NO: 193

Trình tự nucleotit của PDISP/HA B Florida với vòng phân giải protein loại bớt (cấu trúc 2102 và 2104)

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGATCGAA
TCTGCACTGGAATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCCACTCAAGGGGAGGTCAATGTGACTGGTGT
GATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGCCCA
GACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACCAATGTGTGTGGGGACCACACCTTCGGCGAAGGCTT
CAATACTCCACGAAGTCAAACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAAAATCAGGCAACTACC
CAATCTTCTCAGAGGATATGAAAATATCAGGCTATCAACCCAAAACGTCATCGATGCGGAAAAGGCACCAGGAGGACCC
TACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAGAGCGGATTTTTCGCAACAATGGCTTGGGCTGTCC
CAAAGGACAACAACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAAAT
CACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATGGAGACTCAAATCCTCAAAAGTTCACC
TCATCTGCTAATGGAGTAACCACACACTATGTTTCTCAGATTGGCAGCTTCCCAGATCAAACAGAAGACGGAGGACTAC
CACAAAGCGGCAGGATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTACCAAAGAGGTGT
TTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCTTTAATTGGTGAAGCA
GATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAA
ATTGCCCAATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAGG
AATGATTGCAGGCTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCGGACCTTAAGAGTACGCAA
GAAGCTATAAACAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAATCTTCAAAGACTAAGTGGTG
CCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCGCA
AATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTGGCACTTGAGAGAAAA
CTAAAGAAAATGCTGGGTCCCTCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAACACAAGTGCAACCAGACCT
GCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCACTGAACATTACTGC
TGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTTCTAGTTTGGCTGTA
ACATTGATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCCATCTGTCTATAA

Fig 61E SEQ ID NO: 194

Trình tự axit amin của PDISP/HA B Florida với vòng phân giải protein loại bớt (cấu trúc 2102 và 2104)

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTRTRGKLCP
DCLNCTDLDVALGRPMCVGTTPSAKASILHEVKPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDAEKAPGGP
YRLGTSGSCPNATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMKNLYGDSNPQKFT
SSANGVTTHYVSQIGSFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEA

DCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQ
EAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERK
LKKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNHTILLYYSTAASSLAV
TLMLAIFIVYMVSRDNVSCSICL*

Fig 61F SEQ ID NO: 195

Cát xét biểu hiện số 2102 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. PDISP/HA B Florida với trình tự nucleotit vòng phân giải protein loại bớt được gạch dưới.

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTG
AGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGAT
AGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGAC
AGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAG
TGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCA
AAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTT
ATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATG
CCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTC
TTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCT
TCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAA
ACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGA
TCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTG
GACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCT
GGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCT
TGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGA
AACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCC
AATACCGCGGAGAAAATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGA
TCTTCGCGGATCGAATCTGCACTGGAATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCCACTCAAGGGGAGGT
CAATGTGACTGGTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGA
GGGAAACTATGCCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACCAATGTGTGTGGGGACCACAC
CTTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAA
AATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAAATATCAGGCTATCAACCCAAAACGTCATCGATGCGGAAAAG
GCACCAGGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAGAGCGGATTTTTCGCAACAA
TGGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGA
AGGGGAAGACCAAATCACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATGGAGACTCAAAT
CCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACACTATGTTTCTCAGATTGGCAGCTTCCCAGATCAAACAG
AAGACGGAGGACTACCACAAAGCGGCAGGATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGT
CTACCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCT
TTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATG
CAAAAGCCATAGGAAATTGCCCAATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAAATATAGACCTCCTGG
TGGAGGATGGGAAGGAATGATTGCAGGCTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCGGAC
CTTAAGAGTACGCAAGAAGCTATAAACAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAATCTTC
AAAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGA
CACTATAAGCTCGCAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTG
GCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAACACA
AGTGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTC
ACTGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCT
TCTAGTTTGGCTGTAACATTGATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCCATCT
GTCTATAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTG
TGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACA
CAAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCG
TTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTG
AATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGC
AATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTAT
GTTACTAGAT

Fig 61G

Biểu diễn bằng sơ đồ của cấu trúc số 2102

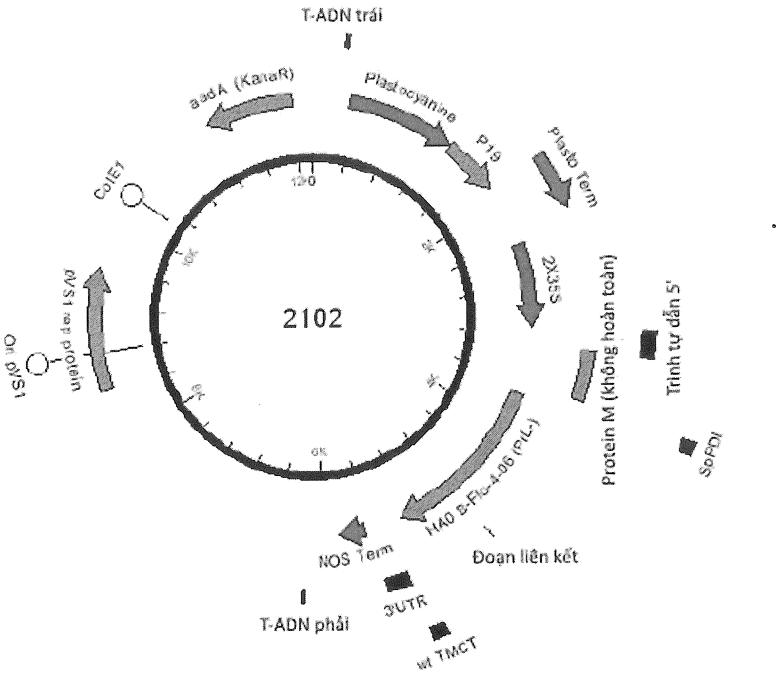


Fig 61H SEQ ID NO: 196

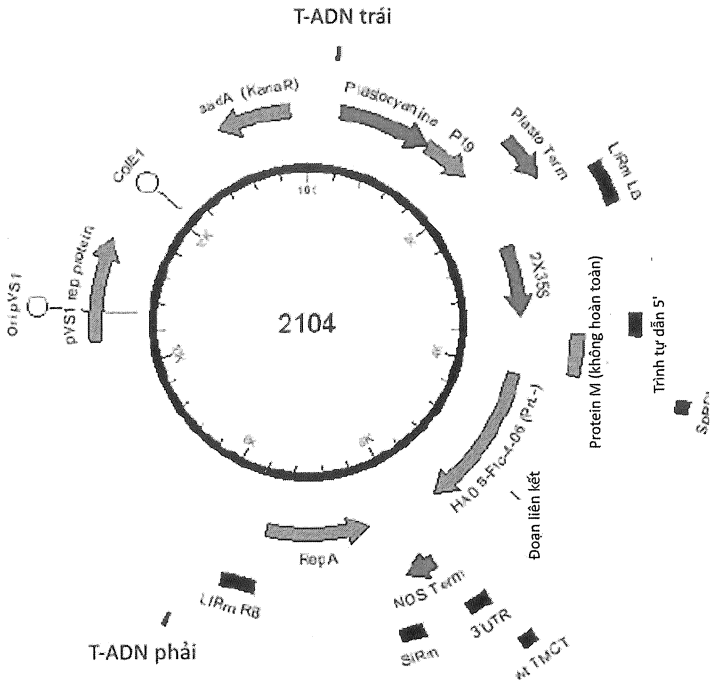
Cát xét biểu hiện số 2104 từ bên trái sang bên phải LIR. PDISP/HA B Florida với trình tự nucleotit vòng phân giải protein loại bớt được gạch dưới.

TGTTGTTGTGACTCCGAGGGGTTGCCTCAAACTCTATCTTATAACCGGCGTGGAGGCATGGAGGCAGGGGTATTTTGGT
CATTTTAATAGATAGTGGAAAATGACGTGGAATTTACTTAAAGACGAAGTCTTTGCGACAAGGGGGGGCCCACGCCGAA
TTTAATATTACCGGCGTGGCCCCCCCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAATTTCCC
GCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGATGGTATTTGGTCGACAAGCTTGCATGCCGG
TCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGA
GACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATA
GTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACA
GTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGT
GGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTA
TTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGC
CTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCT
TCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTT
CCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAAA
CCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGAT
CTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGG
ACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTG
GAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTT
GACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAA
ACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCA
ATACCGCGGAGAAAATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGAT
CTTCGCGGATCGAATCTGCACTGGAATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCCACTCAAGGGGAGGTC
AATGTGACTGGTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAG
GGAAACTATGCCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACCAATGTGTGTGGGGACCACACC
TTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAAA

ATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAAATATCAGGCTATCAACCCAAAACGTCATCGATGCGGAAAAGG
CACCAGGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAGAGCGGATTTTTCGCAACAAT
GGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAA
GGGGAAGACCAAATCACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATGGAGACTCAAATC
CTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACACTATGTTTCTCAGATTGGCAGCTTCCCAGATCAAACAGA
AGACGGAGGACTACCACAAAGCGGCAGGATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTC
TACCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCTT
TAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGC
AAAAGCCATAGGAAATTGCCCAATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAAATATAGACCTCCTGGT
GGAGGATGGGAAGGAATGATTGCAGGCTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCGGACC
TTAAGAGTACGCAAGAAGCTATAAACAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAATCTTCA
AAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGAC
ACTATAAGCTCGCAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTGG
CACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAACACAA
GTGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCA
CTGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTT
CTAGTTTGGCTGTAACATTGATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCCATCTG
TCTATAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGT
GCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACAC
AAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGT
TCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGA
ATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCA
ATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATG
TTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCCATAAAATGATTATTTTATGAATATATTTCATTGTGCAAGTAGAT
AGAAATTACATATGTTACATAACACACGAAATAAACAAAAAAAGACAATCCAAAAACAAACACCCCAAAAAAAATAATC
ACTTTAGATAAACTCGTATGAGGAGAGGCACGTTCAGTGACTCGACGATTCCCGAGCAAAAAAAGTCTCCCCGTCACAC
ATATAGTGGGTGACGCAATTATCTTTAAAGTAATCCTTCTGTTGACTTGTCATTGATAACATCCAGTCTTCGTCAGGAT
TGCAAAGAATTATAGAAGGGATCCCACCTTTTATTTTCTTCTTTTTTCCATATTTAGGGTTGACAGTGAAATCAGACTG
GCAACCTATTAATTGCTTCCACAATGGGACGAACTTGAAGGGGATGTCGTCGATGATATTATAGGTGGCGTGTTCATCG
TAGTTGGTGAAATCGATGGTACCGTTCCAATAGTTGTGTCGTCCGAGACTTCTAGCCCAGGTGGTCTTTCCGGTACGAG
TTGGTCCGCAGATGTAGAGGCTGGGGTGTCGGATTCCATTCCTTCCATTGTCCTGGTTAAATCGGCCATCCATTCAAGG
TCAGATTGAGCTTGTTGGTATGAGACAGGATGTATGTAAGTATAAGCGTCTATGCTTACATGGTATAGATGGGTTTCCC
TCCAGGAGTGTAGATCTTCGTGGCAGCGAAGATCTGATTCTGTGAAGGGCGACACATACGGTTCAGGTTGTGGAGGGAA
TAATTTGTTGGCTGAATATTCCAGCCATTGAAGTTTTGTTGCCCATTCATGAGGGAATTCTTCCTTGATCATGTCAAGA
TATTCCTCCTTAGACGTTGCAGTCTGGATAATAGTTCTCCATCGTGCGTCAGATTTGCGAGGAGAGACCTTATGATCTC
GGAAATCTCCTCTGGTTTTAATATCTCCGTCCTTTGATATGTAATCAAGGACTTGTTTAGAGTTTCTAGCTGGCTGGAT
ATTAGGGTGATTTCCTTCAAAATCGAAAAAAGAAGGATCCCTAATACAAGGTTTTTTATCAAGCTGGAGAAGAGCATGA
TAGTGGGTAGTGCCATCTTGATGAAGCTCAGAAGCAACACCAAGGAAGAAAATAAGAAAAGGTGTGAGTTTCTCCCAGA
GAAACTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGATAGGTAAGGAAAACATATTTAGATTGGAGTCTGAAGTT
CTTACTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACTCTATCTTATAACCGGCGTGGAGGCATGGA
GGCAGGGGTATTTTGGTCATTTTAATAGATAGTGGAAAATGACGTGGAATTTACTTAAAGACGAAGTCTTTGCGACAAG
GGGGGGCCCACGCCGAATTTAATATTACCGGCGTGGCCCCCCCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTT
AAAGTAGAAAATTTCCCGCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGATGGTATTTG

Fig 61I

Biểu diễn bằng sơ đồ của cấu trúc số 2104



2X35S/CPMV HT+/PDISP/B Florida+H1 California TMCT với vòng phân giải protein loại bớt/NOS (Cấu trúc số 2016) và 2X35S/CPMV HT+/BeYDV/PDISP/B Florida+H1 California TMCT với vòng phân giải protein loại bớt/NOS (Cấu trúc số 2108)

FIG 62A SEQ ID NO: 197

IF-H1cTMCT.S1-4r (Cấu trúc 2106 và 2108)

ACTAAAGAAAATAGGCCTTTAAATACATATTCTACACTGTAGAGAC

FIG 62B SEQ ID NO: 198

Trình tự nucleotit của PDISP/HA B Florida+H1Cal TMCT với vòng phân giải protein loại bớt (cấu trúc số 2106 và 2108)

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCGGATCGAA
TCTGCACTGGAATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCCACTCAAGGGGAGGTCAATGTGACTGGTGT
GATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAGGGAAACTATGCCCA
GACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACCAATGTGTGTGGGGACCACACCTTCGGCGAAGGCTT
CAATACTCCACGAAGTCAAACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAAAATCAGGCAACTACC
CAATCTTCTCAGAGGATATGAAAATATCAGGCTATCAACCCAAAACGTCATCGATGCGGAAAAGGCACCAGGAGGACCC
TACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAGAGCGGATTTTTCGCAACAATGGCTTGGGCTGTCC
CAAAGGACAACAACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAAAT
CACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATGGAGACTCAAATCCTCAAAAGTTCACC
TCATCTGCTAATGGAGTAACCACACACTATGTTTCTCAGATTGGCAGCTTCCCAGATCAAACAGAAGACGGAGGACTAC
CACAAAGCGGCAGGATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTACCAAAGAGGTGT
TTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCTTTAATTGGTGAAGCA
GATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAA
ATTGCCCAATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAAATATAGACCTCCTGGTGGAGGATGGGAAGG
AATGATTGCAGGCTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCGGACCTTAAGAGTACGCAA

GAAGCTATAAACAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAATCTTCAAAGACTAAGTGGTG
CCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCGCA
AATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTGGCACTTGAGAGAAAA
CTAAAGAAAATGCTGGGTCCCTCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAACACAAGTGCAACCAGACCT
GCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCACTGAACATTACTGC
TGCATCTTTAAATGATGATGGATTGGATAACTACCAGATTTTGGCGATCTATTCAACTGTCGCCAGTTCATTGGTACTG
GTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAATATGTATTTAA

FIG 62C SEQ ID NO: 199

Trình tự axit amin của PDISP/HA B Florida+H1Cal TMCT với vòng phân giải protein loại bớt (cấu trúc số 2106 và 2108)

MAKNVAIFGLLFSLLVLVPSQIFADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSYFANLKGTRTRGKLCP
DCLNCTDLDVALGRPMCVGTTPSAKASILHEVKPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSTQNVIDAEKAPGGP
YRLGTSGSCPNATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICTEGEDQITVWGFHSDNKTQMKNLYGDSNPQKFT
SSANGVTTHYVSQIGSFPDQTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSLPLIGEA
DCLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKTPLKLANGTKYRPPGGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQ
EAINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERK
LKKMLGPSAVEIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNITAASLNDDGLDNYQILAIYSTVASSLVL
VVSLGAISFWMCSNGSLQCRICI*

FIG 62D SEQ ID NO: 200

Cát xét biểu hiện số 2106 từ đoạn khởi đầu 2X35S đến đoạn kết thúc NOS. PDISP/HA B Florida+H1Cal TMCT với trình tự nucleotit vòng phân giải protein loại bớt được gạch dưới.

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTG
AGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGAT
AGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGAC
AGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAG
TGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCA
AAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTT
ATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATG
CCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTC
TTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCT
TCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAA
ACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGA
TCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTG
GACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCT
GGAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCT
TGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGA
AACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCC
AATACCGCGGAGAAAATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGA
TCTTCGCGGATCGAATCTGCACTGGAATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCCACTCAAGGGGAGGT
CAATGTGACTGGTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGA
GGGAAACTATGCCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACCAATGTGTGTGGGGACCACAC
CTTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAA
AATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAAATATCAGGCTATCAACCCAAAACGTCATCGATGCGGAAAAG
GCACCAGGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAGAGCGGATTTTTCGCAACAA
TGGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGA
AGGGGAAGACCAAATCACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATGGAGACTCAAAT
CCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACACTATGTTTCTCAGATTGGCAGCTTCCCAGATCAAACAG
AAGACGGAGGACTACCACAAAGCGGCAGGATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGT
CTACCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCT
TTAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATG

CAAAAGCCATAGGAAATTGCCCAATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAAATATAGACCTCCTGG
TGGAGGATGGGAAGGAATGATTGCAGGCTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCGGAC
CTTAAGAGTACGCAAGAAGCTATAAACAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAATCTTC
AAAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGA
CACTATAAGCTCGCAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTG
GCACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAACACA
AGTGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTC
ACTGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAACTACCAGATTTTGGCGATCTATTCAACTGTCGCC
AGTTCATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAATAT
GTATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTG
TGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACA
CAAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCG
TTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTG
AATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGC
AATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTAT
GTTACTAGAT

FIG 62E

Biểu diễn bằng sơ đồ của cấu trúc 2106

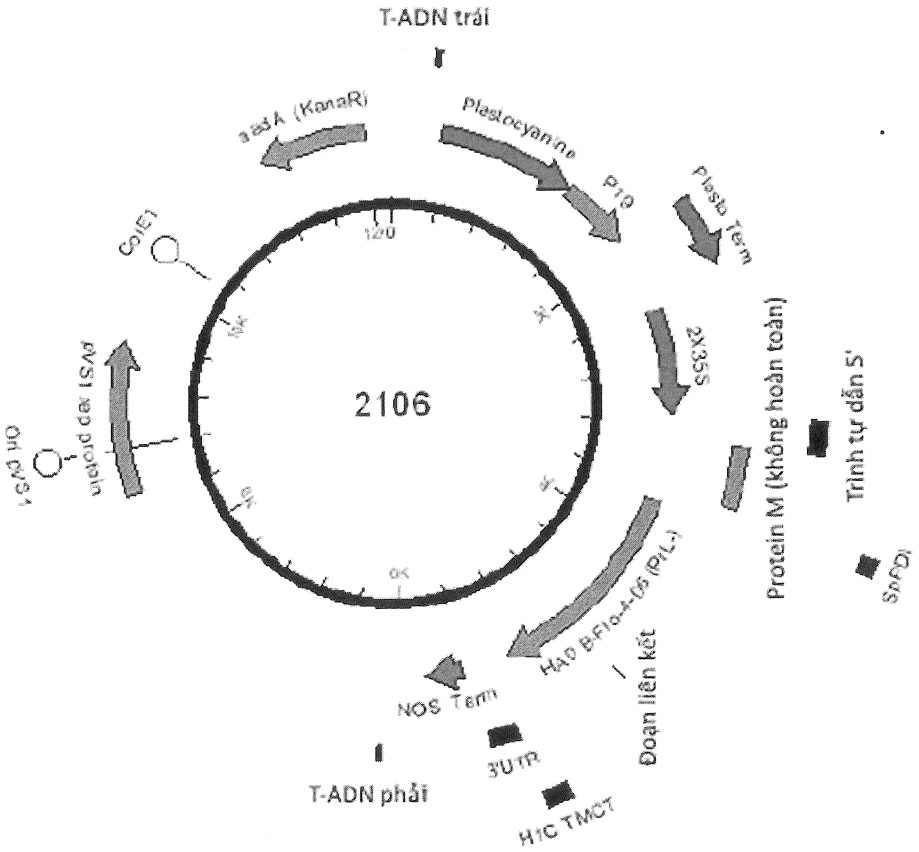


FIG 62F SEQ ID NO: 201

Cát xét biểu hiện số 2108 từ bên trái sang bên phải LIR. PDISP/HA B Florida+H1Cal TMCT với trình tự nucleotit vòng phân giải protein loại bớt được gạch dưới.

TGTTGTTGTGACTCCGAGGGGTTGCCTCAAACTCTATCTTATAACCGGCGTGGAGGCATGGAGGCAGGGGTATTTTGGT
CATTTTAATAGATAGTGGAAAATGACGTGGAATTTACTTAAAGACGAAGTCTTTGCGACAAGGGGGGGCCCACGCCGAA
TTTAATATTACCGGCGTGGCCCCCCCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAATTTCCC
GCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGATGGTATTTGGTCGACAAGCTTGCATGCCGG
TCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGA
GACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATA
GTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACA
GTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGT
GGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTA
TTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGC
CTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCT
TCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTT
CCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAAA
CCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTCTTGTCTTTCTTGCGTGAGCGAT
CTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTGG
ACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCCTATTCTTGTCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTG
GAGGCTGCTGTTCAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTT
GACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAA
ACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCA
ATACCGCGGAGAAAATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGAT
CTTCGCGGATCGAATCTGCACTGGAATAACATCTTCAAACTCACCTCATGTGGTCAAAACAGCCACTCAAGGGGAGGTC
AATGTGACTGGTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAGGACCAGAG
GGAAACTATGCCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACCAATGTGTGTGGGGACCACACC
TTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTACATCCGGGTGCTTTCCTATAATGCACGACAGAACAAAA
ATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAAATATCAGGCTATCAACCCAAAACGTCATCGATGCGGAAAAGG
CACCAGGAGGACCCTACAGACTTGGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAGAGCGGATTTTTCGCAACAAT
GGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAA
GGGGAAGACCAAATCACTGTTTGGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATGGAGACTCAAATC
CTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACACTATGTTTCTCAGATTGGCAGCTTCCCAGATCAAACAGA
AGACGGAGGACTACCACAAAGCGGCAGGATTGTTGTTGATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTC
TACCAAAGAGGTGTTTTGTTGCCTCAAAAGGTGTGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCTT
TAATTGGTGAAGCAGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGC
AAAAGCCATAGGAAATTGCCCAATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAAATATAGACCTCCTGGT
GGAGGATGGGAAGGAATGATTGCAGGCTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCGGACC
TTAAGAGTACGCAAGAAGCTATAAACAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAAGAATCTTCA
AAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGAC
ACTATAAGCTCGCAAATAGAACTTGCAGTCTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTGG
CACTTGAGAGAAAACTAAAGAAAATGCTGGGTCCCTCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAACACAA
GTGCAACCAGACCTGCTTAGACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCA
CTGAACATTACTGCTGCATCTTTAAATGATGATGGATTGGATAACTACCAGATTTTGGCGATCTATTCAACTGTCGCCA
GTTCATTGGTACTGGTAGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGCTCTAATGGGTCTCTACAGTGTAGAATATG
TATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGT
GCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACAC
AAAAAGATTTTAATTTTATTAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGT
TCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGA
ATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCA
ATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATG
TTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCCATAAAATGATTATTTTATGAATATATTTCATTGTGCAAGTAGAT
AGAAATTACATATGTTACATAACACACGAAATAAACAAAAAAAGACAATCCAAAAACAAACACCCCAAAAAAAATAATC

ACTTTAGATAAACTCGTATGAGGAGAGGCACGTTCAGTGACTCGACGATTCCCGAGCAAAAAAAGTCTCCCCGTCACAC
ATATAGTGGGTGACGCAATTATCTTTAAAGTAATCCTTCTGTTGACTTGTCATTGATAACATCCAGTCTTCGTCAGGAT
TGCAAAGAATTATAGAAGGGATCCCACCTTTTATTTTCTTCTTTTTTCCATATTTAGGGTTGACAGTGAAATCAGACTG
GCAACCTATTAATTGCTTCCACAATGGGACGAACTTGAAGGGGATGTCGTCGATGATATTATAGGTGGCGTGTTCATCG
TAGTTGGTGAAATCGATGGTACCGTTCCAATAGTTGTGTCGTCCGAGACTTCTAGCCCAGGTGGTCTTTCCGGTACGAG
TTGGTCCGCAGATGTAGAGGCTGGGGTGTCGGATTCCATTCCTTCCATTGTCCTGGTTAAATCGGCCATCCATTCAAGG
TCAGATTGAGCTTGTTGGTATGAGACAGGATGTATGTAAGTATAAGCGTCTATGCTTACATGGTATAGATGGGTTTCCC
TCCAGGAGTGTAGATCTTCGTGGCAGCGAAGATCTGATTCTGTGAAGGGCGACACATACGGTTCAGGTTGTGGAGGGAA
TAATTTGTTGGCTGAATATTCCAGCCATTGAAGTTTTGTTGCCCATTCATGAGGGAATTCTTCCTTGATCATGTCAAGA
TATTCCTCCTTAGACGTTGCAGTCTGGATAATAGTTCTCCATCGTGCGTCAGATTTGCGAGGAGAGACCTTATGATCTC
GGAAATCTCCTCTGGTTTTAATATCTCCGTCCTTTGATATGTAATCAAGGACTTGTTTAGAGTTTCTAGCTGGCTGGAT
ATTAGGGTGATTTCCTTCAAAATCGAAAAAAGAAGGATCCCTAATACAAGGTTTTTTATCAAGCTGGAGAAGAGCATGA
TAGTGGGTAGTGCCATCTTGATGAAGCTCAGAAGCAACACCAAGGAAGAAAATAAGAAAAGGTGTGAGTTTCTCCCAGA
GAAACTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGATAGGTAAGGAAAACATATTTAGATTGGAGTCTGAAGTT
CTTACTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACTCTATCTTATAACCGGCGTGGAGGCATGGA
GGCAGGGGTATTTTGGTCATTTTAATAGATAGTGGAAAATGACGTGGAATTTACTTAAAGACGAAGTCTTTGCGACAAG
GGGGGGCCCACGCCGAATTTAATATTACCGGCGTGGCCCCCCCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTT
AAAGTAGAAAATTTCCCGCCCACTAGGGTTAAAGGTGTTCACACTATAAAAGCATATACGATGTGATGGTATTTG

FIG 62G

Biểu diễn bằng sơ đồ của cấu trúc số 2108

