



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

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

(11)



1-0023591

(51)⁷ C07K 14/11; A61K 39/145; A61P
31/16; C12N 7/01; C12N 15/44; C12N
15/82; C12N 7/00; A01H 5/00; C07K
19/00

(13) B

(21) 1-2012-00186

(22) 25/06/2010

(86) PCT/CA2010/000983 25/06/2010

(87) WO2010/148511 29/12/2010

(30) 61/220,161 24/06/2009 US

(45) 27/04/2020 385

(43) 25/07/2012 292A

(73) MEDICAGO INC. (CA)

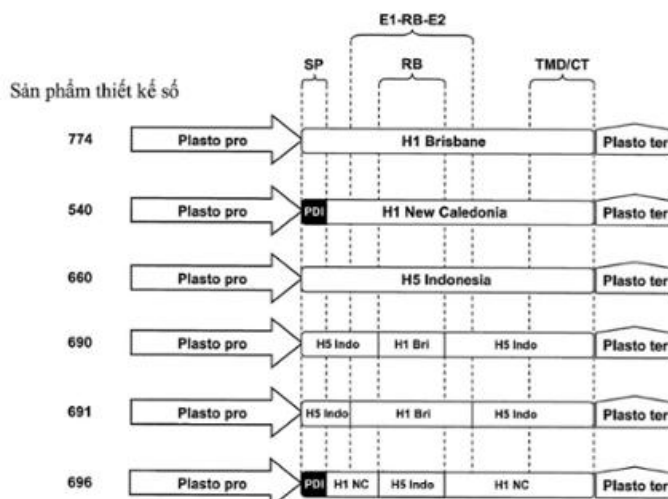
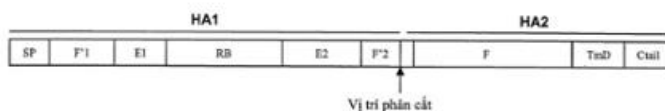
1020 route de l'Église, bureau 600, Québec (Québec) G1V 3V9, Canada

(72) COUTURE, Manon (CA); DARGIS, Michele (CA); LAVOIE, Pierre-Olivier (CA);
VEZINA, Louis-Philippe (CA); D'AOUST, Marc-Andre (CA)

(74) Công ty Luật TNHH AMBYS Hà Nội (AMBYS HANOI)

(54) PHƯƠNG PHÁP SẢN XUẤT HẠT TƯƠNG TỰ VIRUT CÚM THỂ KHẢM Ở THỰC VẬT VÀ AXIT NUCLEIC MÃ HÓA POLYPEPTIT CỦA HẠT TƯƠNG TỰ VIRUT NÀY

(57) Sáng chế đề cập đến phương pháp tổng hợp các hạt tương tự virus cúm (VLP) ở thực vật hoặc bộ phận của thực vật. Phương pháp này gồm bước biểu hiện HA cúm thể khảm trong cây hoặc bộ phận của cây. Sáng chế cũng đề cập đến VLP chứa protein cúm HA thể khảm và các lipid thực vật. Sáng chế còn đề cập đến axit nucleic mã hóa HA cúm thể khảm cũng như các vectơ. Các VLP có thể được sử dụng để bào chế các vắc xin cúm, hoặc có thể được sử dụng để nâng cao chất lượng các vắc xin hiện nay.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các hạt tương tự virus. Cụ thể hơn, sáng chế đề cập đến các hạt tương tự virus thể khảm chứa chất ngưng kết hồng cầu cúm, và các phương pháp sản xuất các hạt tương tự virus cúm thể khảm.

Tình trạng kỹ thuật của sáng chế

Cúm là nguyên nhân dẫn đến tử vong ở người do virus hô hấp, và trong đợt “cúm mùa”, ước lượng có khoảng từ 10 đến 20% dân số trên toàn thế giới có thể bị nhiễm, dẫn đến từ 250 đến 500000 người chết mỗi năm.

Phương pháp phổ biến chống lại bệnh cúm ở người là tiêm chủng hàng năm. Vaccin thường là sự kết hợp của các chủng khác nhau mà chúng được dự đoán là những chủng trội dẫn tới cúm theo mùa, tuy nhiên số lượng liều vaccin được sản xuất hàng năm không đủ để tiêm chủng cho toàn bộ dân số thế giới. Ví dụ, Canada và Hoa Kỳ đạt đủ số liều vaccin để chủng ngừa cho khoảng một phần ba dân số của họ, trong khi chỉ có 17% dân số của Châu Âu có thể được chủng ngừa - sự sản xuất này không đủ để đối mặt với đại dịch cúm trên thế giới. Dù là việc sản xuất cần thiết hàng năm có thể bằng cách này hay cách khác gặp trong năm nhất định, các chủng chiếm ưu thế thay đổi từ năm này sang năm khác, do đó sự dự trữ nhiều lần với nhu cầu thấp trong năm là không thực tế. Về phương diện kinh tế, việc sản xuất quy mô lớn vaccin cúm có hiệu lực là mối quan tâm lớn của chính phủ và cũng như công nghiệp tư nhân.

Glycoprotein bề mặt chất ngưng kết hồng cầu cúm (HA) là cả protein liên kết thụ thể và protein dung hợp màng. Đây là trimer của các nhóm con giống hệt, mỗi nhóm chứa hai polypeptit liên kết disulfua, HA1 và HA2, nhận được từ sự phân cắt phân giải protein của tiền chất, HA0, có trình tự peptit tín hiệu tại đầu tận cùng N của nó và trình tự dạng móc của màng tại đầu tận cùng C của nó. Sự phân cắt tạo thành HA1 và HA2 tạo ra đầu tận cùng N của polypeptit nhỏ hơn, HA2, có trình tự dạng móc của màng tại đầu tận cùng C của nó. Sự phân cắt cần đến hoạt tính dung hợp màng nhưng không phải để tạo miễn dịch. Trình tự đầu tận cùng N HA2 được gọi là “peptit dung hợp” do sự phân cắt tại các trình tự

kị nước tương tự cũng cần đến hoạt tính của các protein dung hợp virus khác, và do peptit tổng hợp 20-chuỗi tận tương tự với các trình tự dung hợp màng *in vitro*.

Nói chung, bề mặt của ‘đầu’ có dạng hình cầu bao gồm các vòng linh hoạt khác nhau có các vùng kháng nguyên biến thiên và đặc trưng thích hợp được thiết kế dưới dạng các vị trí A, B, C, D và E (được đưa ra trong Wiley et al., 1987. *Annu. Rev Biochem* 56:365-394). Sự xen vào hoặc thay thế các trình tự peptit ngắn tại một vài vị trí (ví dụ B và E) trong các nghiên cứu miễn dịch được mô tả (Garcia-Sastré et al. 1995. *Biologicals* 23:171-178). Nhân tố sinh trưởng biểu bì (EGF), kháng thể chuỗi đơn (scFV) và vùng Fc của IgG, có kích thước nằm trong khoảng từ 53 đến 246 axit amin, được xen vào tại đầu tận cùng N của H7 và các thể khảm được biểu hiện thành công (Hatzioannou et al., 1999. *Human Gene Therapy* 10:1533-1544). Gần đây, 90 và 140 các vùng axit amin của kháng nguyên bảo vệ *Bacillus anthracis* được dung hợp vào các đầu tận cùng amino của H3 (Li et al., 2005. *J. Virol* 79:10003-10002). Copeland (Copeland et al., 2005. *J. Virol* 79:6459-6471) mô tả sự biểu hiện glycoprotein bề mặt gp120 Env HIV trên thân H3, tại đó vùng gp120 được thay thế toàn bộ đầu dạng hình cầu của HA.

Các sản phẩm tái tổ hợp khác nhau được phát triển dưới dạng các đối tượng vaccin cúm tái tổ hợp. Các phương pháp này tập trung biểu hiện, sản xuất, và tinh sạch các protein HA và NA cúm typ A, gồm biểu hiện các protein này sử dụng các tế bào côn trùng bị nhiễm baculovirus (Crawford et al, 1999 *Vaccine* 17:2265-74; Johansson, 1999 *Vaccine* 17:2073-80), các vectơ của virus, và các cấu trúc vaccin ADN (Olsen et al., 1997 *Vaccine* 15:1149-56).

Việc sản xuất chủng virus cúm không lây nhiễm với các mục đích vaccin là một cách để tránh sự nhiễm không cố ý. Mặt khác, các hạt tương tự virus (các VLP) dưới dạng các phần tử thể cho virus đã nuôi cấy được nghiên cứu. Các VLP bắt chước cấu tạo của vỏ capsid virus, nhưng thiếu hệ gen, và do đó không thể tái tạo hoặc tạo ra các biện pháp nhiễm khuẩn thứ phát. Các kỹ thuật sản xuất VLP cúm hiện nay dựa trên sự biểu hiện đồng thời của nhiều protein virus, và sự phụ thuộc này thể hiện nhược điểm của các kỹ thuật này từ trường hợp đại dịch và bệnh dịch hàng năm, thời gian đáp ứng quyết định sự tiêm chủng. Một hệ thống sản xuất VLP đơn giản hơn, ví dụ, hệ thống dựa vào sự biểu hiện của chỉ một hoặc một vài protein của virus mà không cần phải biểu hiện các protein của virus không cấu trúc là mong muốn để thúc đẩy sự phát triển các vaccin.

Các virus có vỏ bọc có thể thu được bao lipid của chúng khi ‘sinh chồi’ mà không có tế bào bị nhiễm và thu được màng từ màng tế bào, hoặc từ tế bào của bào quan bên trong. Trong các hệ tế bào động vật có vú và baculovirus, ví dụ, virus cúm sinh chồi từ màng tế bào (Quan et al 2007 J. Virol 81:3514-3524). Chỉ một vài virus có vỏ bọc được biết tới làm lây nhiễm các cây (ví dụ, các chi của Tospovirus và Rhabdovirus). Các virus có vỏ bọc của cây đã biết, chúng được mô tả nhờ sinh chồi từ màng bên trong của tế bào chủ, và không từ màng tế bào. Mặc dù số lượng nhỏ các VLP tái tổ hợp được sản xuất trong các vật chủ của cây, nhưng không một VLP nào có nguồn gốc từ màng tế bào, điều này đưa ra câu hỏi là liệu các VLP có nguồn gốc từ màng tế bào, bao gồm các VLP cúm có thể được tạo ra trong cây không.

Sự tạo thành các VLP, trong hệ thống bất kỳ, các vị trí quan trọng đòi hỏi cấu trúc của các protein – thay đổi các đoạn trình tự ngắn tương ứng với các vòng bề mặt được chọn của cấu trúc có dạng hình cầu có thể không có nhiều hiệu quả biểu hiện polypeptit của chính nó, tuy nhiên các nghiên cứu về cấu trúc không chứng minh được hiệu quả của các biến đổi dựa trên sự tạo thành các VLP. Sự kết hợp của các vùng và cấu trúc khác nhau của HA (ví dụ trình tự dạng móc của màng, các vùng thân hoặc gốc của trimer tách riêng khỏi đầu hình cầu từ màng) được tiến hóa bởi virus và không cải thiện được các biến đổi tương tự mà không làm mất tính nguyên vẹn của HA trimer và sự tạo thành VLP. Sự sản xuất các HA VLP cúm được mô tả trước đây bởi các tác giả sáng chế trong WO 2009/009876.

Bản chất kỹ thuật của sáng chế

Sáng chế đề cập đến các hạt tương tự virus. Cụ thể hơn, sáng chế đề cập đến các hạt tương tự virus chứa chất ngưng kết hồng cầu cúm thể khảm, và các phương pháp sản xuất các hạt tương tự virus chất ngưng kết hồng cầu cúm thể khảm.

Mục tiêu của sáng chế là đề xuất các hạt tương tự virus cúm thể khảm được cải thiện (VLP).

Sáng chế đề xuất polypeptit bao gồm HA cúm thể khảm bao gồm nhóm vùng gốc (SDC), nhóm vùng đầu (HDC) và nhóm vùng xuyên màng (TDC) trong đó: SDC chứa vùng phụ F'1, F'2 và F; HDC chứa vùng phụ RB, E1 và E2; TDC bao gồm vùng phụ TmD và Ctail; và trong đó ít nhất một vùng phụ là HA cúm thứ nhất và các vùng phụ khác trong một hoặc nhiều HA cúm thứ hai. HA cúm thứ nhất và thứ hai có thể độc lập

được chọn từ nhóm bao gồm H1, H3, H5 và B. Hơn nữa, polypeptit có thể bao gồm peptit tín hiệu.

Sáng chế cũng đề xuất axit nucleic mã hóa polypeptit bao gồm HA cúm thể khảm bao gồm nhóm vùng gốc (SDC), nhóm vùng đầu (HDC) và nhóm vùng xuyên màng (TDC) trong đó: SDC chứa vùng phụ F'1, F'2 và F; HDC chứa vùng phụ RB, E1 và E2; TDC bao gồm vùng phụ TmD và Ctail; và trong đó ít nhất một vùng phụ là HA cúm thứ nhất, và các vùng phụ khác trong một hoặc nhiều HA cúm thứ hai. Axit nucleic cũng có thể mã hóa polypeptit bao gồm peptit tín hiệu cộng thêm SDC, HDC và TDC như được định rõ.

Phương pháp sản xuất các hạt tương tự virus cúm thể khảm (các VLP) ở thực vật cũng được đề xuất, phương pháp bao gồm:

a) tạo ra axit nucleic mã hóa HA cúm thể khảm peptit tín hiệu, nhóm vùng gốc (SDC), nhóm vùng đầu (HDC) và nhóm vùng xuyên màng (TDC) trong đó: SDC chứa vùng phụ F'1, F'2 và F; HDC chứa vùng phụ RB, E1 và E2; TDC chứa vùng phụ TmD và Ctail; và trong đó ít nhất một vùng phụ là HA cúm thứ nhất, và các vùng phụ khác trong một hoặc nhiều HA cúm thứ hai vào cây, hoặc bộ phận của cây, và

b) ủ cây, hoặc bộ phận của cây, dưới các điều kiện cho phép biểu hiện axit nucleic, nhờ đó sản xuất được các VLP.

Sáng chế gồm phương pháp được mô tả ở trên, trong đó trong bước đưa vào (bước a), axit nucleic được đưa vào cây theo kiểu tạm thời. Mặt khác, trong bước đưa vào (bước a), axit nucleic được đưa vào cây và được hợp nhất ổn định. Phương pháp có thể bao gồm bước: c) thu vật chủ và tinh sạch các VLP.

Sáng chế đề xuất cây, hoặc bộ phận của cây bao gồm HA cúm thể khảm, hoặc trình tự nucleotit mã hóa HA cúm thể khảm, HA cúm thể khảm bao gồm nhóm vùng gốc (SDC), nhóm vùng đầu (HDC) và nhóm vùng xuyên màng (TDC) trong đó: SDC chứa vùng phụ F'1, F'2 và F; HDC chứa vùng phụ RB, E1 và E2; TDC chứa vùng phụ TmD và Ctail; và trong đó ít nhất một vùng phụ là HA cúm thứ nhất và các vùng phụ khác trong một hoặc nhiều HA cúm thứ hai.

Cây, hoặc bộ phận của cây, có thể còn bao gồm axit nucleic bao gồm trình tự nucleotit mã hóa một hoặc nhiều hơn một protein chaperon được liên kết có hiệu quả với

vùng điều hòa có hoạt tính trong cây. Một hoặc nhiều hơn một protein chaperon có thể được chọn từ nhóm gồm Hsp40 và Hsp70.

Sáng chế gắn liền với hạt tương tự virus (VLP) bao gồm HA cúm thể khả, HA cúm thể khả bao gồm nhóm vùng gốc (SDC), nhóm vùng đầu (HDC) và nhóm vùng xuyên màng (TDC) trong đó: SDC chứa vùng phụ F'1, F'2 và F; HDC chứa vùng phụ RB, E1 và E2; TDC chứa vùng phụ TmD và Ctail; và trong đó ít nhất một vùng phụ là HA cúm thứ nhất và các vùng phụ khác trong một hoặc nhiều HA cúm thứ hai. VLP có thể bao gồm thêm các N-glycan đặc trưng ở thực vật, hoặc N-glycan được cải biến.

Chế phẩm bao gồm liều dùng hữu hiệu VLP như được mô tả và chất mang được dụng cũng được đề xuất.

Theo một khía cạnh khác của sáng chế đề xuất phương pháp gây ra sự miễn dịch với sự nhiễm virus cúm trong đối tượng, bao gồm việc sử dụng VLP cho đối tượng. VLP có thể được sử dụng cho đối tượng theo đường miệng, trong da, trong mũi, trong cơ, trong bụng, trong tĩnh mạch, hoặc dưới da.

Các vùng điều hòa có thể được gắn hiệu quả với trình tự mã hóa protein HA thể khả gồm các vùng điều hòa có hoạt tính trong tế bào thực vật, tế bào côn trùng hoặc tế bào nấm men. Các vùng điều hòa này có thể gồm vùng điều hòa plastocyanin, vùng điều hòa Ribulosa 1,5-bisphosphat carboxylaza/oxygenaza (RuBisCO), protein gắn chlorophyll a/b (CAB) hoặc ST-LS1. Các vùng điều hòa khác gồm 5' UTR, 3' UTR hoặc các trình tự kết thúc. Vùng điều hòa plastocyanin có thể là vùng điều hòa plastocyanin ở cỏ linh lăng; 5' UTR, 3'UTR hoặc các trình tự kết thúc cũng có thể là trình tự cỏ linh lăng.

Sáng chế đề xuất polypeptit HA cúm thể khả bao gồm bệnh cúm thứ nhất và bệnh cúm thứ hai, bệnh cúm thứ nhất và bệnh cúm thứ hai có thể độc lập được chọn từ nhóm bao gồm B, H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15 và H16; với điều kiện là bệnh cúm thứ nhất và bệnh cúm thứ hai không thuộc typ cúm, typ phụ, hoặc có nguồn gốc tương tự.

Phù hợp với một vài khía cạnh của sáng chế, polypeptit HA cúm thể khả bao gồm trình tự peptit tín hiệu, trình tự peptit tín hiệu có thể được chọn từ nhóm bao gồm trình tự peptit tín hiệu gốc, trình tự peptit tín hiệu PDI cỏ linh lăng, trình tự peptit tín hiệu H5 cúm và trình tự peptit tín hiệu H1 cúm

Sáng chế đề xuất phương pháp sản xuất VLP chứa chất ngưng kết hồng cầu (HA) cúm thể khảm trong vật chủ có khả năng sản xuất VLP, gồm cây, côn trùng, hoặc nấm men bao gồm, sự đưa vào axit nucleic mã hóa HA cúm thể khảm bao gồm bao gồm nhóm vùng gốc (SDC), nhóm vùng đầu (HDC) và nhóm vùng xuyên màng (TDC) trong đó: SDC chứa vùng phụ F'1, F'2 và F; HDC chứa vùng phụ RB, E1 và E2; TDC chứa vùng phụ TmD và Ctail; và trong đó ít nhất một vùng phụ là HA cúm thứ nhất và các vùng phụ khác trong một hoặc nhiều HA cúm thứ hai, vào vật chủ, và ủ vật chủ trong các điều kiện cho phép biểu hiện axit nucleic, nhờ đó sản xuất được các VLP.

Sản xuất các VLP trong cây thể hiện các ưu điểm khác so với sản xuất các hạt này trong canh trường tế bào côn trùng. Các lipid của cây có thể kích thích các tế bào miễn dịch đặc hiệu và làm tăng đáp ứng miễn dịch gây ra. Các màng của cây được tạo thành bởi lipid, phosphatidylcholin (PC) và phosphatidyletanolamin (PE), và cũng chứa các glycosphingolipit chỉ liên quan đến cây và một số vi khuẩn và động vật nguyên sinh. Sphingolipit khác thường trong đó không phải là este của glyxerol tương tự PC hoặc PE nhưng đúng hơn gồm rượu amino mạch dài tạo thành liên kết amit với mạch axit béo chứa nhiều hơn 18 cacbon. PC và PE cũng như các glycosphingolipit có thể gắn với các phân tử CD1 được biểu hiện bởi các tế bào miễn dịch của động vật có vú như các tế bào trình diện kháng nguyên (các APC) giống các tế bào tua và đại thực bào và các tế bào khác gồm các tế bào lympho B và T trong tuyến ức và gan. Hơn nữa, ngoài tác dụng có hiệu quả của tá được khi có mặt lipid thực vật, thì khả năng của các N-glycan thực vật tạo thuận lợi để bắt giữ các kháng nguyên glycoprotein bởi các tế bào trình diện kháng nguyên, có thể là ưu điểm để sản xuất các VLP thể khảm ở thực vật. Không mong muốn bị ràng buộc bởi lý thuyết, người ta dự đoán rằng các VLP thể khảm làm từ thực vật gây ra phản ứng miễn dịch mạnh hơn so với các VLP thể khảm được tạo thành trong các hệ thống sản xuất khác và phản ứng miễn dịch này được gây ra bởi các VLP thể khảm làm từ thực vật này sẽ mạnh hơn khi được so sánh với phản ứng miễn dịch được gây ra bởi các vaccin virus giảm động lực nguyên vẹn hoặc sống.

Trái với các vaccin được làm từ toàn bộ virus, các VLP thể khảm tạo ra ưu điểm khi chúng không có tính lây nhiễm, do đó biện pháp ngăn chặn sinh học giới hạn không còn là vấn đề quan trọng vì vaccin được làm từ virus nguyên vẹn, lây nhiễm, và nó không cần để sản xuất. Các VLP thể khảm làm từ cây tạo ra ưu điểm cho phép hệ thống biểu hiện được

phát triển trong nhà kính hoặc trên cánh đồng, do đó kinh tế hơn đáng kể và thích hợp với quy mô lớn.

Ngoài ra, các cây không bao gồm các enzym có liên quan đến sự tổng hợp và bổ sung các gốc axit sialic vào protein. Các VLP có thể được tạo ra khi không có neuraminidaza (NA), và không cần biểu hiện đồng thời NA, hoặc xử lý các tế bào sản xuất hoặc chiết với sialidaza (neuraminidaza), để đảm bảo sự sản xuất VLP ở thực vật.

Bản tóm tắt sáng chế không nhất thiết phải mô tả tất cả các đặc trưng của sáng chế.

Mô tả vắn tắt các hình vẽ

Đặc trưng này và các đặc trưng khác của sáng chế sẽ trở nên rõ ràng hơn từ phần mô tả sau đây trong đó tiến hành tham khảo các hình vẽ đính kèm trong đó:

Fig.1A thể hiện biểu đồ của các vùng phụ HA. SP: peptit tín hiệu, F'1, F'2, và F: vùng phụ dung hợp; RB: vùng phụ liên kết thụ thể, E1 và E2: vùng phụ esteraza, TMD/CT: vùng phụ đuôi tế bào chất và màng vận chuyển. Fig.1B chỉ ra sự thể hiện dưới dạng biểu đồ của các catxet biểu hiện dựa trên plastoxyanin (số sản phẩm thiết kế: 774, 540, 660, 690, 691, 696) để biểu hiện chất ngưng kết hồng cầu H1 A/Brisbane/59/2007 (H1/Bri), chất ngưng kết hồng cầu H1 A/New Caledonia/20/99 (H1/NC) và chất ngưng kết hồng cầu H5 A/Indonesia/5/05 (H5/Indo) có dạng gốc và dạng thể khảm. Plasto pro: trình tự khởi đầu plastoxyanin ở cỏ linh lăng, Plasto ter: nhân tố kết thúc plastoxyanin ở cỏ linh lăng, SP: peptit tín hiệu, RB: vùng phụ liên kết thụ thể, E1-RB-E2: esteraza và vùng phụ liên kết thụ thể, TMD/CT: vùng phụ đoạn cuối tế bào chất và màng vận chuyển, PDI: protein disulfua isomeraza ở cỏ linh lăng. Fig.1C chỉ ra sự liên kết trình tự axit amin được chồng lên bởi sự liên kết về cấu trúc với các HA cúm khác nhau (B/Florida/4/2006 (B/Florida), SEQ ID NO:94 (ngân hàng gen số truy cập ACA33493.1; B/Malaysia/2506/2004 (B-Malaysia), SEQ ID NO:95 (ngân hàng gen số truy cập ABU99194.1; H1/Bri (A-Brisbane), SEQ ID NO:96 (ngân hàng gen số truy cập ADE28750.1; H1 A/Solomon Islands/3/2006 (A-Sol.Isl), SEQ ID NO:97 (ngân hàng gen số truy cập ABU99109.1); H1/NC (A-NewCal), SEQ ID NO:98 (ngân hàng gen số truy cập AAP34324.1; H2 A/Singapore/1/1957 (A-Singapore), SEQ ID NO:99 (ngân hàng gen số truy cập AAA64366.1); H3 A/Brisbane/10/2007 (A-Brisbane), SEQ ID NO:100 (ngân hàng gen số truy cập ACI26318.1); H3 A/Wisconsin/67/2005(A-WCN), SEQ ID NO:101 (ngân hàng gen số truy cập ABO37599.1); H5 A/Anhui/1/2005 (A-Anhui), SEQ ID

NO:102 (ngân hàng gen số truy cập ABD28180.1); H5 A/Vietnam/1194/2004(A-Vietnam), SEQ ID NO:103 (ngân hàng gen số truy cập ACR48874.1); H5-Indo, SEQ ID NO:104 (ngân hàng gen số truy cập ABW06108.1. Ranh giới giữa F'1, Esteraza 1, liên kết thụ thể, Esteraza 2, F'2, Dung hợp peptit, các vùng phụ TMD/CT, và các cầu di-sulfua được chỉ ra.

Fig.2 thể hiện trình tự axit amin của các vùng phụ được chỉ ra của HA thể khảm được biểu hiện bởi, khung phía trên, các sản phẩm thiết kế 690, 734 (SEQ ID NO: 11), 696 (SEQ ID NO: 112), và khung phía dưới, 691 (SEQ ID NO: 113. Axit amin 1-92 của SEQ ID NO: 111 là vùng F'1+E1 của H5/Indo; các axit amin 93-263 là vùng đầu RB của H1/Brisbane và axit amin 264-552 là vùng E2+F'2 của H5/Indo. Axit amin 1-92 của SEQ ID NO: 112 là vùng F'1+E1 của H5/NC; các axit amin 93-301 là vùng đầu RB của H5/Indo và axit amin 302-586 là vùng E2+F'2 của H1/NC. Axit amin 1-42 của SEQ ID NO: 113 là vùng F'1+E1 của H5/Indo; các axit amin 43-273 là vùng đầu E1-RB-E2 của H1/Brisbane và axit amin 274-552 là vùng F'2 của H5/Indo.

Fig.3 thể hiện trình tự axit amin của vùng mã hóa các sản phẩm thiết kế 690 và 734 (SEQ ID NO: 80) bao gồm vùng phụ RB của H1/Bri, peptit tín hiệu H5/Indo, và phức hệ vùng gốc (SDC) bao gồm các vùng phụ H5/Indo F'1, E1, E2, F'2 và F.

Fig.4 thể hiện trình tự axit amin của vùng mã hóa sản phẩm thiết kế 691 (SEQ ID NO: 81) bao gồm phức hệ vùng đầu H1/Bri (HDC) bao gồm peptit tín hiệu E1, RB, E2, H5/Indo, và phức hệ vùng gốc H5/Indo (SDC) bao gồm vùng phụ H5/Indo F'1, F'2 và F.

Fig.5 thể hiện trình tự axit amin của vùng mã hóa sản phẩm thiết kế 696 (SEQ ID NO: 82) bao gồm vùng phụ RB của H5/Indo, peptit tín hiệu PDI, và phức hệ vùng gốc H1/NC bao gồm F'1, E1, E2 và F'2.

Fig.6 thể hiện kỹ thuật phân tích thẩm tách miễn dịch biểu hiện H1/Bri dưới dạng gốc, sản phẩm thiết kế 774 (bao gồm H1/Bri), sản phẩm thiết kế 692 (bao gồm phức hệ vùng đầu (HDC) của H1/Bri), và sản phẩm thiết kế 690 (bao gồm vùng phụ RB của H1/Bri được dung hợp bởi phức hệ vùng gốc H5/Indo (SDC) ở thực vật. Với mỗi sản phẩm thiết kế, tổng số protein chiết từ 3 cây riêng biệt được phân tích. Hai mươi microgam protein được nạp cho mỗi cây được phân tích. Thẩm tách Western được bộc lộ bởi các kháng thể đơn dòng kháng HA (kháng H1-Brisbane; FII 10-I50). Sản phẩm thiết

kế 774 biểu hiện H1/Bri với peptit tín hiệu gốc của H1/Bri; sản phẩm thiết kế 690, 691 biểu hiện HA bởi peptit tín hiệu gốc của H5/Indo.

Fig.7 thể hiện kỹ thuật phân tích thẩm tách miễn dịch của H5/Indo có dạng gốc, sản phẩm thiết kế 660 (bao gồm H5/Indo, hoặc sản phẩm thiết kế 696 (bao gồm vùng phụ H1/Indo RB được dung hợp bởi H1/NC SDC, các vùng phụ E1 và E2). Với mỗi sản phẩm thiết kế, tổng số protein chiết từ 3 cây riêng biệt được phân tích. Hai mươi microgam protein được nạp cho mỗi cây được phân tích. Thẩm tách Western được bộc lộ bởi các kháng thể đa dòng kháng H5 Indonesia (ITC IT-003-005V). Sản phẩm thiết kế 660 biểu hiện H5/Indo với peptit tín hiệu gốc của nó; sản phẩm thiết kế 696 biểu hiện HA thể khảm với peptit tín hiệu PDI.

Fig.8 thể hiện sự biểu hiện dưới dạng biểu đồ các catxet biểu hiện dựa trên 35SCPMV/HT để biểu hiện H1/Bri có dạng gốc (sản phẩm thiết kế 732) và dạng thể khảm (các sản phẩm thiết kế 733 và 734). Sản phẩm thiết kế 733, bao gồm peptit tín hiệu PDI và HDC, SDC và phức hệ vùng xuyên màng (TDC) của H1/Bri, và sản phẩm thiết kế 734 bao gồm peptit tín hiệu H5/Indo, F'1, E1, E2, F'2, F, và RB từ H1/Bri. 35S pro: trình tự khởi đầu phiên mã CaMV 35S, NOS ter: nhân tố kết thúc nopalín synthaza, SP: peptit tín hiệu, RB: vùng phụ liên kết thụ thể, E1-RB-E2: esteraza và vùng phụ liên kết thụ thể, TMD/CT: vùng phụ đoạn cuối tế bào chất và màng vận chuyển, PDI: protein disulfua isomerase ở cơ nhộng; CPMV-HT: các yếu tố 5' và 3' trong hệ thống biểu hiện virus khảm đậu đũa có khả năng dịch mã cao.

Fig.9 thể hiện kỹ thuật thẩm miễn dịch biểu hiện H1/Bri có dạng gốc, sản phẩm thiết kế 732 (bao gồm H1/Bri dưới sự kiểm soát catxet biểu hiện dựa trên 35SCPMV/HT), sản phẩm thiết kế 733 (bao gồm peptit tín hiệu PDI được dung hợp với H1/Bri; nhờ sự kiểm soát catxet biểu hiện dựa trên 35SCPMV/HT), hoặc sản phẩm thiết kế 734 bao gồm vùng phụ H1/Bri RB được dung hợp với các vùng phụ H5/Indo SDC, E1 và E2; nhờ sự kiểm soát catxet biểu hiện dựa trên 35SCPMV/HT). Với mỗi sản phẩm thiết kế, tổng số protein chiết từ 3 cây riêng biệt được phân tích. Năm microgam protein được nạp cho mỗi cây được phân tích. Thẩm tách Western được bộc lộ bởi các kháng thể đơn dòng kháng HA (FII 10-I50).

Fig.10 thể hiện sự thể hiện dưới dạng biểu đồ các catxet biểu hiện dựa trên 35SCPMV/HT để biểu hiện chất ngưng kết hồng cầu H3 A/Brisbane/10/2007 HA

(H3/Bri) và B/Florida/4/2006 HA (B/Flo). Sản phẩm thiết kế 736 bao gồm H3/Bri được dung hợp với peptit tín hiệu PDI. Sản phẩm thiết kế 737 bao gồm H3/Bri được dung hợp với peptit tín hiệu PDI và H5/Indo TMD/CT. Sản phẩm thiết kế 739 bao gồm B/Flo được dung hợp với peptit tín hiệu PDI. Sản phẩm thiết kế 745 bao gồm B/Flo được dung hợp với peptit tín hiệu PDI và H5/Indo TMD/CT. 35S pro: trình tự khởi đầu phiên mã CaMV 35S, NOS ter: nhân tố kết thúc nopal synthase, SP: peptit tín hiệu, RB: vùng phụ liên kết thụ thể, E1-RB-E2: esterase và vùng phụ liên kết thụ thể, TMD/CT: vùng phụ đoạn cuối tế bào chất và màng vận chuyển, PDI: protein disulfide isomerase ở cơ linh lăng; CPMV-HT: các yếu tố 5' và 3' trong hệ thống biểu hiện Virut khảm đậu đũa có khả năng dịch mã cao.

Fig.11 thể hiện biên dung hợp trong các sản phẩm thiết kế số 745 và 737. Nguồn gốc của trình tự HA được chỉ ra bởi các mũi tên đầu tròn. Các axit amin trong vùng xuyên màng là QILSIYSTVA, và được đặt trước bởi axit amin là một phần của vùng ngoài.

Fig.12 thể hiện trình tự axit amin của chất ngưng kết hồng cầu H5/H3 thể khảm (SEQ ID NO: 83; sản phẩm thiết kế 737) bao gồm peptit tín hiệu PDI, vùng ngoài của H3 A/Brisbane/10/2007 và TMD/CT của H5 A/Indonesia/5/2005.

Fig.13 thể hiện trình tự axit amin của chất ngưng kết hồng cầu H5/B thể khảm (SEQ ID NO: 84) bao gồm vùng ngoài của B/Florida/4/2006 và TMD/CT của H5 A/Indonesia/5/2005 được mã hóa bằng khung đọc mở trong sản phẩm thiết kế số 745.

Fig.14 thể hiện sự phân tích thẩm tách miễn dịch sự biểu hiện của B/Flo dạng gốc, sản phẩm thiết kế 739 (bao gồm PDI-B/Flo), hoặc với sản phẩm thiết kế 745 (bao gồm B/Flo HDC và SDC được dung hợp với H5/Indo TDC). Với mỗi sản phẩm thiết kế, tổng số protein chiết từ 3 cây riêng biệt được phân tích. Hai mươi microgam protein được nạp cho mỗi cây được phân tích. Thẩm tách Western được bộc lộ bởi các kháng thể đa dòng kháng HA B/Florida (NIBSC 07/356).

Fig.15 thể hiện phân tích thẩm tách miễn dịch sự biểu hiện của H3/Bri có dạng gốc, sản phẩm thiết kế 736 (bao gồm PDI sp - H3/Bri), hoặc với sản phẩm thiết kế 737 (bao gồm H3/Bri HDC và SDC được dung hợp với H5/Indo TDC). Với mỗi sản phẩm thiết kế, tổng số protein chiết từ 3 cây riêng biệt được phân tích. Hai mươi microgam protein được nạp cho mỗi cây được phân tích. Thẩm tách Western được bộc lộ bởi các kháng thể đa dòng kháng H3 Brisbane (NIBSC 08/124).

Fig.16 thể hiện phương pháp sắc ký loại trừ kích thước của protein từ lá chiết từ các cây được thăm bởi sản phẩm thiết kế số 745. Hàm lượng protein tương đối của các phần rửa giải được thể hiện cho mỗi phần. Kỹ thuật thẩm tách Western chất ngưng kết hồng cầu sử dụng các kháng thể đa dòng kháng HA B/Florida (NIBSC 07/356) trong các phần từ 7 đến 15 được thể hiện nhờ đồ thị. Đỉnh rửa giải của Blue Dextran 2000 được chỉ ra bởi mũi tên (phần số 8).

FIG.17 thể hiện trình tự axit nucleic (SEQ ID NO: 52) của đoạn được tổng hợp bao gồm vùng mã hóa H5 hoàn chỉnh (A/Indonesia/5/05 (H5N1)) (gồm peptit tín hiệu và codon kết thúc) được chặn, tại đầu 5' bởi vị trí HindIII và, tại đầu 3', bởi vị trí SacI.

FIG.18 thể hiện trình tự axit nucleic (SEQ ID NO: 53) của sản phẩm thiết kế 660, và catxet biểu hiện HA bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cở linh lăng và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu có dạng H5 A/Indonesia/5/05 (H5N1), plastoxyanin ở cở linh lăng 3' UTR và trình tự kết thúc.

FIG.19 thể hiện trình tự axit nucleic (SEQ ID NO: 54) của trình tự mã hóa H1 loại đại (A/New Caledonia/20/99 (H1N1) (Ngân hàng gen số truy cập AY289929) mà không có TmD và Ctail.

FIG.20 thể hiện trình tự axit nucleic (SEQ ID NO: 55) của đoạn tổng hợp bao gồm trình tự mã hóa H1 (A/New Caledonia/20/99 (H1N1) không có TmD và Ctail. Trong vùng 5', các nucleotit cuối cùng có bắt nguồn từ PDI SP và gồm vị trí giới hạn BglII và tại đầu 3', vị trí SacI/StuI kép trực tiếp tìm thấy cùng chiều với codon kết thúc.

FIG.21 thể hiện trình tự axit nucleic (SEQ ID NO: 56) của đoạn được tổng hợp bao gồm trình tự mã hóa C-ter H1 (A/New Caledonia/20/99 (H1N1) gồm TmD và Ctail từ vị trí KpnI đến codon kết thúc (bị chặn tại đầu 3' nhờ vị trí SacI/StuI kép).

FIG.22 thể hiện trình tự nucleotit cho *Medicago sativa* mARN đối với protein disulfua isomeraza. Truy cập Ngân hàng gen Số.Z11499 (SEQ ID NO: 57). Các nucleotit 32-103 mã hóa peptit tín hiệu PDI.

Fig.23 thể hiện trình tự nucleotit với plasmit PromPlasto-PDISP-Plasto 3'UTR. FIG.23A thể hiện trình tự nucleotit với PromPlasto-PDISP (SEQ ID NO: 58). FIG.23B thể hiện trình tự nucleotit từ Plasto 3'UTR (SEQ ID NO:85). Peptit tín hiệu protein

disulfua isomeraza (PDI) được gạch chân. Các vị trí giới hạn BglII (AGATCT) và SacI (GAGCTC) được sử dụng để tách dòng được thể hiện bằng cách in đậm.

FIG.24 thể hiện trình tự axit nucleic (SEQ ID NO: 59; sản phẩm thiết kế 540) của catxet biểu hiện HA bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cổ linh lăng và 5' UTR, trình tự mã hóa peptit tín hiệu từ PDI và dạng H1 A/New Caledonia/20/99 (H1N1), plastoxyanin ở cổ linh lăng 3' UTR và trình tự kết thúc. H1 từ trình tự mã hóa A/New Caledonia/20/1999 được gạch chân.

FIG.25 thể hiện trình tự axit nucleic (SEQ ID NO: 60) của đoạn được tổng hợp bao gồm vùng mã hóa H1 hoàn chỉnh (A/Brisbane/59/07 (H1N1)) (gồm peptit tín hiệu và codon kết thúc) bị chặn, tại đầu 5' bởi trình tự gen plastoxyanin ở cổ linh lăng tương ứng với 84 nucleotit thứ nhất ngược chiều với ATG ban đầu, bắt đầu với vị trí DraIII và, tại đầu 3', bởi vị trí SacI.

FIG.26 thể hiện trình tự axit nucleic (SEQ ID NO: 61; sản phẩm thiết kế 774) của catxet biểu hiện HA bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cổ linh lăng và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu của dạng H1 A/Brisbane/59/07 (H1N1), plastoxyanin ở cổ linh lăng 3' UTR và trình tự kết thúc.

Fig.27 thể hiện trình tự axit nucleic của catxet biểu hiện số 828 (SEQ ID NO: 62), từ PacI (ngược chiều với trình tự khởi đầu phiên mã) đến AscI (ngay cùng chiều với nhân tố kết thúc NOS). Máy giải trình tự CPMV HT 3' UTR được gạch dưới bởi ATG đột biến in đậm. Vị trí giới hạn ApaI được gạch dưới và in nghiêng.

Fig.28 thể hiện trình tự axit nucleic (SEQ ID NO: 63; sản phẩm thiết kế 690) của catxet biểu hiện H5/H1 thể khảm bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cổ linh lăng và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, plastoxyanin ở cổ linh lăng 3' UTR và trình tự kết thúc. Trình tự mã hóa HA thể khảm được gạch dưới.

FIG.29 thể hiện trình tự axit nucleic (SEQ ID NO: 64; sản phẩm thiết kế 691) của catxet biểu hiện H5/H1 thể khảm bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cổ linh lăng và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, plastoxyanin ở cổ linh lăng 3' UTR và trình tự kết thúc. Trình tự mã hóa HA thể khảm được gạch dưới.

FIG.30 thể hiện trình tự axit nucleic (SEQ ID NO: 65; sản phẩm thiết kế 696) của catxet biểu hiện H5/H1 thể khảm bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cổ

linh lãg và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, plastoxyanin ở cở linh lãg 3' UTR và trình tự kết thúc. Trình tự mã hóa HA thể khảm được gạch dưới.

FIG.31 thể hiện trình tự axit nucleic (SEQ ID NO: 66; sản phẩm thiết kế 732) của catxet biểu hiện HA bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu có dạng H1 A/Brisbane/59/07 (H1N1), CPMV-HT 3' UTR và trình tự kết thúc NOS. Trình tự mã hóa H1/Bri được gạch dưới.

FIG.32 thể hiện trình tự axit nucleic (SEQ ID NO: 67) của trình tự mã hóa, từ ATG đến kết thúc, của sản phẩm thiết kế trung gian số 787.

FIG.33 thể hiện trình tự axit nucleic (SEQ ID NO: 68; sản phẩm thiết kế số 733) của catxet biểu hiện SpPDI H1/Bri bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa peptit tín hiệu từ PDI, trình tự mã hóa chất ngưng kết hồng cầu có dạng H1 A/Brisbane/59/07 (H1N1), CPMV-HT 3' UTR và trình tự kết thúc NOS. Trình tự mã hóa của SpPDI H1/Bri được gạch dưới.

FIG.34 thể hiện trình tự axit nucleic (SEQ ID NO: 69; sản phẩm thiết kế 734) của catxet biểu hiện H5/H1 thể khảm bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, CPMV-HT 3' UTR và trình tự kết thúc NOS. Trình tự mã hóa HA thể khảm được gạch dưới.

FIG.35 thể hiện trình tự axit nucleic (SEQ ID NO: 70) của đoạn được tổng hợp bao gồm vùng mã hóa H3 hoàn chỉnh (A/Brisbane/10/07 (H3N2)) (gồm peptit tín hiệu và codon kết thúc) bị chặn, tại đầu 5' bởi trình tự gen plastoxyanin ở cở linh lãg tương ứng với 84 nucleotit thứ nhất ngược chiều của ATG ban đầu, bắt đầu với vị trí DraIII và, tại đầu 3', bởi vị trí SacI.

FIG.36 thể hiện trình tự axit nucleic (SEQ ID NO: 71; sản phẩm thiết kế số 736) của catxet biểu hiện HA bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa peptit tín hiệu từ PDI, trình tự mã hóa chất ngưng kết hồng cầu có dạng H3 A/Brisbane/10/07 (H2N3), CPMV-HT 3' UTR và trình tự kết thúc NOS. Trình tự mã hóa của Sp PDI H3/Bris được gạch dưới.

FIG.37 thể hiện trình tự axit nucleic (SEQ ID NO: 72; sản phẩm thiết kế 737) của catxet biểu hiện H5/H3 thể khảm bao gồm trình tự khởi đầu phiên mã CaMV 35S,

CPMV-HT 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, CPMV-HT 3' UTR và trình tự kết thúc NOS. Trình tự mã hóa HA thể khảm được gạch dưới.

FIG.38 thể hiện trình tự axit nucleic (SEQ ID NO: 73) của đoạn được tổng hợp bao gồm vùng mã hóa HA hoàn chỉnh (B/Florida/4/06) (gồm peptit tín hiệu và codon kết thúc) bị chặn, tại đầu 5' bởi trình tự gen plastoxyanin ở cở linh lăng tương ứng với 84 nucleotit thứ nhất ngược chiều của ATG ban đầu, bắt đầu với vị trí DraIII và, tại đầu 3', bởi vị trí SacI.

FIG.39 thể hiện trình tự axit nucleic (SEQ ID NO: 74; sản phẩm thiết kế số 739) của catxet biểu hiện HA bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa peptit tín hiệu từ PDI, trình tự mã hóa chất ngưng kết hồng cầu có dạng HA B/Florida/4/06, CPMV-HT 3' UTR và trình tự kết thúc NOS. Trình tự mã hóa Sp PDI B/Flo được gạch dưới.

FIG.40 thể hiện trình tự axit nucleic (SEQ ID NO: 75; sản phẩm thiết kế 745) của catxet biểu hiện H5/B thể khảm bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, CPMV-HT 3' UTR và trình tự kết thúc NOS. Trình tự mã hóa HA thể khảm được gạch dưới.

FIG.41 thể hiện trình tự axit nucleic mã hóa Msj1 (SEQ ID NO: 76).

FIG.42 thể hiện trình tự axit nucleic (SEQ ID NO: 77) của một phần sản phẩm thiết kế số R850, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu phiên mã) tới EcoRI (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa HSP40 được gạch dưới.

FIG.43 thể hiện trình tự axit nucleic (SEQ ID NO: 78) của một phần sản phẩm thiết kế số R860, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu phiên mã) tới EcoRI (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa HSP70 được gạch dưới.

FIG.44 thể hiện trình tự axit nucleic (SEQ ID NO: 79) của một phần sản phẩm thiết kế số R870, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu phiên mã) tới EcoRI (ngay ngược chiều với nhân tố kết thúc NOS). Trình tự mã hóa HSP40 được gạch dưới in nghiêng và trình tự mã hóa HSP70 được gạch dưới. A) các nucleotit 1-4946; B) các nucleotit 4947-9493.

Fig.45 thể hiện dưới dạng biểu đồ sản phẩm thiết kế số R472.

Fig.46 chỉ ra mẫu cầu nối disulfua của cúm A. Đánh số cầu nối: 1) Cys4HA1-Cys137HA2, 2) Cys60HA1-Cys72HA1, 3) Cys94HA1-Cys143HA1, 4) Cys292HA1-Cys318HA1 5) Cys144HA2-Cys148HA2 và 6) Cys52HA1-Cys277HA1. Các cầu nối disulfua khác với typ phụ A và B (Fig.47) được chỉ ra bởi các mũi tên. Đánh số từ H3 protein thuần thực được sử dụng.

Fig.47 chỉ ra mẫu cầu nối disulfua của cúm B HA. Đánh số cầu nối: 1) Cys4HA1-Cys137HA2, 2) Cys60HA1-Cys72HA1, 3) Cys94HA1-Cys143HA1, 4) Cys292HA1-Cys318HA1 5) Cys144HA2-Cys148HA2, 6) Cys52HA1-Cys277HA1, 7) Cys54HA1-Cys57HA1 và 8) Cys178HA1-Cys272HA1. Các cầu nối disulfua khác với typ phụ A (Fig.46) và B được chỉ ra bởi các mũi tên. Đánh số từ H3 protein hoàn thiện được sử dụng.

Fig.48 chỉ ra biểu đồ các chức năng dung hợp hoán đổi vùng. Fig.48A chỉ ra sự dung hợp vùng phụ RB từ H1/Bri, H3/Bri, và B/Flo với H5/Indo SDC's, và vùng phụ RB của H5/Indo với vùng gốc H1/NC. Fig.48B chỉ ra sự dung hợp vùng phụ E1-RB-E2 (HDC) từ H1/Bri, H3/Bri hoặc B/Flo với H5/Indo SDC, và H5/Indo HDC với H1/NC SDC.

FIG.49A thể hiện trình tự nucleotit (SEQ ID NO: 86) của H1 A/California/04/09. Trình tự mã hóa peptit tín hiệu protein disulfua isomeraza ở cở linh lăng được gạch dưới và trình tự mã hóa H1 hoàn thiện được làm nổi bật nhờ in đậm. Fig.49B thể hiện trình tự axit amin (SEQ ID NO: 87) của H1 A/California/04/09. Peptit tín hiệu protein disulfua isomeraza ở cở linh lăng được gạch dưới.

Fig.50 thể hiện kỹ thuật phân tích thẩm tách miễn dịch sự biểu hiện chất ngưng kết hồng cầu thể khả H5/B (sản phẩm thiết kế số 747; bao gồm B/Flo HDC và SDC được dung hợp với H5/Indo TDC) sau sự thẩm của AGL1/747 nguyên chất, thẩm đồng thời với AGL1/443 (vector trống) và được thẩm đồng thời với AGL1/R870 (HSP40/HSP70). Với mỗi sản phẩm thiết kế, tổng số protein chiết từ 3 cây riêng biệt được phân tích. Hai mươi microgam protein được nạp cho mỗi cây được phân tích. Thẩm tách Western được bộc lộ bởi các kháng thể đa dòng kháng B-Florida (NIBSC).

FIG.51A thể hiện trình tự nucleotit của trình tự trình tự khởi đầu phiên mã 2X35S (SEQ ID NO:88). FIG.51B thể hiện trình tự nucleotit của sản phẩm thiết kế 747 (SEQ ID NO:93) từ PacI (ngược chiều trình tự khởi đầu phiên mã 35S) tới AscI (nhân tố kết thúc

NOS ngay cùng chiều). Trình tự mã hóa HA thể khảm được gạch dưới. Trình tự trình tự khởi đầu phiên mã 2X35S được chỉ ra nhờ in nghiêng.

Mô tả chi tiết sáng chế

Sáng chế đề cập đến các hạt tương tự virus. Cụ thể hơn, sáng chế đề cập đến các hạt tương tự virus chứa chất ngưng kết hồng cầu cúm thể khảm, và các phương pháp sản xuất các hạt tương tự virus cúm thể khảm.

Sự mô tả dưới đây là một phương án thực hiện được ưu tiên.

Sáng chế đề xuất axit nucleic bao gồm trình tự nucleotit mã hóa chất ngưng kết hồng cầu cúm thể khảm (HA) được liên kết hiệu quả với vùng điều hòa hoạt hóa ở thực vật.

Hơn nữa, sáng chế đề xuất phương pháp sản xuất các hạt tương tự virus (các VLP) ở thực vật. Phương pháp gồm việc đưa axit nucleic mã hóa HA cúm thể khảm được gắn hiệu quả với vùng điều hòa hoạt tính trong cây, vào trong cây, hoặc bộ phận của cây, và ủ cây hoặc bộ phận của cây dưới các điều kiện cho phép biểu hiện axit nucleic, nhờ đó sản xuất ra các VLP.

Sáng chế đề xuất thêm VLP bao gồm HA cúm thể khảm. VLP có thể được sản xuất theo phương pháp được đề xuất bởi sáng chế.

Bởi “protein thể khảm” hoặc “polypeptit thể khảm”, có nghĩa là protein hoặc polypeptit bao gồm các trình tự axit amin từ hai hoặc nhiều hơn hai nguồn, ví dụ nhưng không giới hạn với, hai hoặc nhiều loại hoặc cúm typ phụ, hoặc cúm có nguồn gốc khác nhau, được dung hợp dưới dạng polypeptit đơn. Protein hoặc polypeptit thể khảm có thể gồm peptit tín hiệu tương tự với, hoặc khác loại với, đoạn còn lại của polypeptit hoặc protein. Protein thể khảm hoặc polypeptit thể khảm có thể được tạo ra dưới dạng sản phẩm phiên mã từ trình tự nucleotit thể khảm, và protein thể khảm hoặc polypeptit thể khảm được phân cắt theo sự tổng hợp, và nếu cần, được tập hợp thành protein đa phân. Do đó, protein thể khảm hoặc polypeptit thể khảm cũng gồm protein hoặc polypeptit bao gồm các đơn vị thứ cấp được tập hợp thông qua cầu nối disulfua (tức là protein đa phân). Ví dụ, polypeptit thể khảm bao gồm các trình tự axit amin từ hai hoặc nhiều hơn hai nguồn có thể được xử lý thành các đơn vị thứ cấp, và các đơn vị thứ cấp được tập hợp thông qua các cầu nối disulfua tạo ra protein thể khảm hoặc polypeptit thể khảm (tham khảo các Fig.46 và 47). Polypeptit có thể là chất ngưng kết hồng cầu (HA), và mỗi trình tự trong hai

hoặc nhiều trình tự axit amin cấu tạo nên polypeptit có thể thu được từ HA khác tạo ra HA thể khảm, hoặc HA cúm thể khảm. HA thể khảm cũng gồm trình tự axit amin bao gồm peptit tín hiệu khác loại (protein tiền thân HA thể khảm) được phân cắt sau khi hoặc trong khi tổng hợp protein. Ưu tiên, polypeptit thể khảm, hoặc HA cúm thể khảm không xuất hiện tự nhiên. Axit nucleic mã hóa polypeptit thể khảm có thể được mô tả dưới dạng "axit nucleic thể khảm", hoặc "trình tự nucleotit thể khảm". Hạt tương tự virus gồm có HA thể khảm có thể được mô tả dưới dạng "VLP thể khảm".

HA cúm thể khảm theo các phương án thực hiện khác nhau theo sáng chế có thể bao gồm phức hợp vùng gốc (SDC) phức hợp vùng đầu (HDC) và phức hợp vùng xuyên màng (TDC), tại đó một hoặc nhiều hơn một vùng phụ của SDC, HDC hoặc TDC thuộc typ, typ phụ HA cúm thứ nhất hoặc từ một nguồn gốc, và một hoặc nhiều hơn một vùng phụ của SDC, HDC hoặc TDC từ loại, typ phụ HA cúm thứ hai, hoặc từ nguồn gốc khác hoặc nguồn gốc thứ hai. Như được mô tả ở đây, "SDC" chứa vùng phụ F'1, F'2 và F, "HDC" chứa vùng phụ RB, E1 và E2, "TDC" chứa vùng phụ TmD và Ctail (TMD/CT; tham khảo các hình vẽ Fig.1A, Fig.46 và Fig.47).

Thuật ngữ "hạt tương tự virus" (VLP), hoặc "các hạt tương tự virus" hoặc "các VLP" đề cập đến các cấu trúc mà chúng tự gắn vào và bao gồm các protein có cấu trúc như protein HA cúm. Các VLP và các VLP thể khảm thường có hình thái và tính kháng nguyên giống với các virion được tạo ra trong sự lây nhiễm, nhưng thiếu thông tin di truyền có khả năng tái bản và do đó không có tính lây nhiễm. Các VLP và các VLP thể khảm có thể được tạo ra trong các tế bào vật chủ thích hợp gồm các tế bào chủ thực vật. Sự chiết sau đó từ tế bào chủ và dựa trên sự phân lập và tinh sạch thêm dưới các điều kiện thích hợp, các VLP và các VLP thể khảm có thể được tinh sạch dưới dạng các cấu trúc nguyên vẹn.

Các VLP thể khảm, hoặc các VLP, được tạo ra từ các protein có nguồn gốc từ bệnh cúm, phù hợp với sáng chế không bao gồm protein M1. Protein M1 được biết đến để gắn ARN (Wakefield và Brownlee, 1989) là một tạp chất trong chế phẩm VLP. Sự có mặt của ARN không được mong muốn để thu được sự điều hòa phù hợp với sản phẩm VLP thể khảm, do đó sự bào chế VLP thể khảm không có ARN có thể thuận lợi hơn.

Các VLP thể khảm theo sáng chế có thể được tạo ra trong tế bào chủ được đặc trưng bởi sự thiếu khả năng sialyl hóa các protein, ví dụ tế bào thực vật, tế bào côn trùng, nấm,

và các sinh vật khác gồm bọt biển, động vật ruột khoang, giun đốt, động vật chân đốt, động vật thân mềm, lớp giun tròn, ngành *trochelmintes*, giun dẹp, hàm tơ, động vật xúc tu, vi rút chlamydia, xoắn khuẩn, vi khuẩn gram dương, vi khuẩn cyanobacteria, cổ khuẩn, hoặc các loài tương tự. Tham khảo, ví dụ Gupta et al., 1999. Nucleic Acids Research 27:370-372; Toukach et al., 2007. Nucleic Acids Research 35:D280-D286; Nakahara et al., 2008. Nucleic Acids Research 36:D368-D371. Các VLP thể khảm được sản xuất như được mô tả ở đây thường không bao gồm neuraminidaza (NA). Tuy nhiên, NA có thể được biểu hiện đồng thời với HA nên các VLP bao gồm HA và NA được mong muốn.

Sáng chế cũng đề xuất các VLP bao gồm HA thể khảm thu được vỏ lipit từ màng tế bào trong tế bào trong đó HA thể khảm được biểu hiện. Ví dụ, nếu HA thể khảm được biểu hiện trong hệ dựa vào thực vật, VLP kết quả có thể thu được vỏ lipit từ màng tế bào của tế bào thực vật.

Nói chung, thuật ngữ “lipit” đề cập đến các phân tử có thể hòa tan chất béo (ưa chất béo), các phân tử xuất hiện tự nhiên. VLP thể khảm được tạo ra trong cây theo một số khía cạnh của sáng chế có thể được tạo phức với các lipit có nguồn gốc thực vật. Các lipit có nguồn gốc từ cây có thể có dạng lớp kép lipit, và có thể bao gồm thêm vỏ bao quanh VLP. Các lipit nguồn gốc thực vật có thể bao gồm các thành phần lipit của màng tế bào của cây tại đó VLP được tạo ra, gồm phospholipit, tri-, di- và monoglycerit, cũng như sterol có thể hòa tan chất béo hoặc các chất trao đổi bao gồm sterol. Các ví dụ gồm phosphatidylcholin (PC), phosphatidyletanolamin (PE), phosphatidylinositol, phosphatidylserin, glycosphingolipit, phytosterol hoặc tổ hợp của chúng. Lipit có nguồn gốc từ cây có thể lần lượt được đề cập đến dưới dạng ‘lipit của cây’. Các ví dụ về phytosterol gồm campesterol, stigmasterol, ergosterol, brassicasterol, delta-7-stigmasterol, delta-7-avenasterol, daunosterol, sitosterol, 24-metylcholesterol, cholesterol hoặc beta-sitosterol – tham khảo, ví dụ, Mongrand et al., 2004. Người có trình độ trung bình trong lĩnh vực kỹ thuật này đều hiểu rằng, thành phần lipit trong màng tế bào của tế bào có thể thay đổi cùng với các điều kiện nuôi cấy và phát triển của tế bào hoặc sinh vật, hoặc các loài, từ tế bào thu được. Nói chung, beta-sitosterol là phytosterol phong phú nhất.

Các màng tế bào thường bao gồm các lớp kép lipit, cũng như các protein có các chức năng khác nhau. Các nồng độ được hạn chế của các lipit cụ thể có thể được nhận thấy trong lớp kép lipit, đề cập đến dưới dạng ‘các mảng lipit’. Các vùng siêu nhỏ của mảng lipit này có thể được làm giàu trong các sphingolipit và sterol. Không ràng buộc bởi

lý thuyết, các mảng lipid có thể có vai trò quan trọng trong sự nhập bào và xuất bào, sự đi vào hoặc đi ra của các virus hoặc các tác nhân lây nhiễm khác, sự truyền tính trạng tín hiệu gian bào, sự tương tác với các thành phần cấu trúc khác trong tế bào hoặc sinh vật, như các chất nền nội bào và ngoại bào.

Sáng chế gồm các VLP bao gồm HA thể khảm, của các vùng phụ có thể thu được từ loại, typ phụ bất kỳ của virus cúm có thể nhiễm độc cho người, gồm, ví dụ, các typ hoặc typ phụ B, H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15 và H16. Theo một vài phương án thực hiện, virus cúm có thể thuộc typ hoặc typ phụ H1, H3, H5 hoặc B. Các ví dụ không giới hạn của loại hoặc typ phụ H1, H3, H5 hoặc B gồm typ phụ A/New Caledonia/20/99 (H1N1) (“H1/NC”; SEQ ID NO:56), typ phụ H1 A/California 04/09 (H1N1) (“H1/Cal”; SEQ ID NO:86), typ phụ A/Indonesia/5/05 (H5N1) (“H5/Indo”), A/Brisbane/59/2007 (“H1/Bri”), và B/Florida/4/2006 (“B/Flo”) và H3 A/Brisbane/10/2007 (“H3/Bri”). Ngoài ra, HA thể khảm có thể bao gồm một hoặc nhiều vùng phụ của chất ngưng kết hồng cầu được phân lập từ một hoặc nhiều virus cúm được nhận diện mới hoặc mới xuất hiện.

Sáng chế cũng đề cập tới các virus cúm nhiễm bệnh cho động vật có vú hoặc các động vật chủ khác, ví dụ người, động vật linh trưởng, ngựa, lợn, chim, chim nước, chim di cư, chim cú, vịt, ngan, ngỗng, gia cầm, gà, lạc đà, thú giống chó, chó, thú giống mèo, mèo, hổ, báo, cây hương, chồn vizon, chồn bạc ửng, chồn sương, vật nuôi kiểng, thú nuôi, chuột nhắt, chuột, hải cẩu, cá voi và các loài tương tự. Một vài virus cúm có thể bị nhiễm nhiều hơn một động vật chủ.

Liên quan tới virus cúm, thuật ngữ “chất ngưng kết hồng cầu” hoặc “HA” theo sáng chế đề cập tới glycoprotein có cấu trúc dạng hạt tương tự virus cúm. Cấu trúc của chất ngưng kết hồng cầu cúm đã được nghiên cứu tỉ mỉ và chứng minh được rằng mức bảo tồn cao dạng cấu trúc bậc hai, bậc ba và bậc bốn. Sự bảo tồn cấu trúc này được quan sát cả với trình tự axit amin có thể thay đổi (tham khảo, ví dụ, Skehel và Wiley, 2000 *Ann Rev Biochem* 69:531-69; Vaccaro et al 2005; được kết hợp ở đây để tham khảo). Trình tự nucleotit mã hóa HA cũng đã được biết đến, và sẵn có ví dụ, từ BioDefense và Public Health Database (ví dụ tại URL: biohealthbase.org/GSearch/home.do?decorator=Influenza) hoặc dữ liệu do National Center for Biotechnology Information lưu trữ (NCBI; ví dụ tại địa chỉ URL:

ncbi.nlm.nih.gov/sites/entrez?db=nucleotide&cmd=search&term=influenza) cả hai được kết hợp ở đây để tham khảo.

Monome HA có thể được phân đoạn thành ba vùng chức năng - vùng gốc, đám vùng gốc (SDC), vùng đầu hình cầu, hoặc đám vùng đầu (HDC) và đám vùng xuyên màng (TDC). SDC bao gồm bốn vùng phụ, peptit dung hợp F, F'1 và F'2 (vùng phụ này thường được đề cập đến là 'đoạn chính'). TDC bao gồm hai vùng phụ, màng vận chuyển (TmD) và đầu tận cùng C (CT). HDC bao gồm ba vùng phụ, các vùng esterase tiêu giảm E1' và E2, và vùng liên kết thụ thể RB. SDC và HDC có thể được đề cập đến chung đến dưới dạng 'vùng ngoài'. Công bố bởi Ha et al. 2002 (EMBO J. 21:865-875; được kết hợp ở đây để tham khảo) minh họa sự định hướng tương đối của các vùng phụ khác nhau của SDC và HDC trong các typ phụ cúm khác nhau, dựa trên các cấu trúc tinh thể học Xray. Biểu đồ của các vùng phụ tương ứng với các đầu tận cùng N và C của các polypeptit HA1 và HA2 được chỉ ra trên Fig.1A. Sự sắp xếp cấu trúc được chú thích của các typ phụ cúm khác nhau được tạo ra trên Fig.1C.

Sự thay đổi axit amino được dung nạp trong chất ngưng kết hồng cầu của các virus cúm. Sự thay đổi này cung cấp các chủng mới liên tục được nhận dạng. Tính lây nhiễm giữa các chủng mới có thể thay đổi. Tuy nhiên, sự tạo thành các trimer chất ngưng kết hồng cầu, duy trì sự tạo thành các VLP sau đó. Do đó, sáng chế, đề xuất trình tự axit amin chất ngưng kết hồng cầu bao gồm HA thể khảm, hoặc axit nucleic mã hóa trình tự axit amin chất ngưng kết hồng cầu, tạo thành các VLP trong cây, và gồm các trình tự đã biết và các trình tự HA biến thể có thể phát triển. Sáng chế cũng đề cập đến việc sử dụng polypeptit HA thể khảm bao gồm TDC, SDC và HDC. Ví dụ protein HA thể khảm có thể là HA0, hoặc HA thể khảm được phân cắt bao gồm các vùng phụ của HA1 và HA2 từ hai hoặc nhiều loại cúm. Protein HA thể khảm có thể được sử dụng trong sản xuất hoặc tạo thành các VLP sử dụng thực vật, hoặc tế bào thực vật, hệ thống biểu hiện.

HA0 có thể được biểu hiện và gấp nếp tạo thành trimer, sau đó có thể tập hợp lại thành các VLP. Sự phân cắt HA0 tạo ra các polypeptit HA1 và HA2 được liên kết bởi cầu nối disulfua (tham khảo các Fig.1C, 46 và 47 để minh họa các mẫu cầu nối disulfua). Với hạt tương tự virus lây nhiễm, sự phân cắt tiền chất HA0 cần thiết để khởi động sự thay đổi thể cấu tạo của HA2 giải phóng peptit dung hợp (tại đầu tận cùng N của polypeptit HA2) và làm cho nó có giá trị để dung hợp tế bào và màng của virus. Tuy nhiên, các VLP không lây nhiễm, và sự phân cắt HA thành HA1 và HA2 là không cần thiết, ví dụ, để sản xuất

vacxin. Tiền chất HA0 không được phân cắt cũng gắn trong các trime và sinh chồi từ màng tế bào tạo thành các hạt nano VLP.

Polypeptit HA0 bao gồm các vùng khác nhau. Vùng phụ RB của HDC bao gồm các vòng khác nhau trong các vùng kháng nguyên được định rõ dưới dạng vị trí A-E. Các kháng thể có thể là các virus cúm lây nhiễm trung tính thường xuyên hướng tới một hoặc nhiều vị trí này. Các vùng phụ esterase tiêu giảm (E1 và E2) có thể có vai trò dung hợp, và có thể liên kết với Ca^{++} . Các vùng F, F'1 và F'2 tương tác và hợp tác tạo thành gốc, nâng đầu của HA trime lên trên màng. TmD và CT có thể gồm trong sự tạo móc HA được gấp nếp vào màng. TmD có thể có vai trò với ái lực của HA với các mảng lipid, trong khi CT có thể có vai trò trong sự bài tiết HA, trong khi một vài chuỗi tận xystein được tìm thấy trong vùng phụ CT có thể được được cải biến lipid. Peptit tín hiệu (SP) có thể được tìm thấy tại đầu tận cùng N của polypeptit HA0. Fig.2, và các bảng 4 và 5 đề xuất các ví dụ về các trình tự axit amin của các vùng SP, F'1, F'2, E1, RB, E2 và F của một vài typ phụ virus cúm.

Xử lý trình tự peptit tín hiệu đầu tận cùng N (SP) trong khi biểu hiện và/hoặc tiết ra các chất ngưng kết hồng cầu gây bệnh cúm có thể có vai trò trong quá trình gấp nếp HA. Thuật ngữ “peptit tín hiệu” thường đề cập tới trình tự axit amin ngắn (khoảng từ 5 tới 30 axit amin), thường được tìm thấy tại đầu tận cùng N của polypeptit chất ngưng kết hồng cầu có thể hướng tới sự chuyển đoạn của polypeptit được dịch mã mới sang một bào quan đặc biệt, hoặc hỗ trợ sự định vị các vùng đặc hiệu của chuỗi polypeptit tương ứng với các vùng khác. Peptit tín hiệu của các chất ngưng kết hồng cầu hướng tới sự chuyển đoạn protein vào trong mạng lưới nội chất và được đề xuất để hỗ trợ sự định vị vùng gần đầu tận cùng N tương ứng với vùng dạng móc màng của polypeptit chất ngưng kết hồng cầu mới sinh để hỗ trợ sự phân cắt và gấp nếp chất ngưng kết hồng cầu hoàn chỉnh.

Sự xen HA vào trong màng lưới nội chất (ER) của tế bào chủ, sự phân cắt peptit tín hiệu và glycosyl hóa protein là các trường hợp dịch mã đồng thời. Sự nếp nếp chính xác của HA cần đến sự glycosyl hóa protein và sự hình thành ít nhất 6 liên kết disulfua trong chuỗi (tham khảo Fig.46 và 47). Trên Fig.46, HA từ typ phụ A được chỉ ra có 6 cầu nối disulfua được bảo tồn cho mỗi monome. Bằng cách so sánh, monome B HA (Fig.47) có bảy cầu nối disulfua, và năm trong số các cầu nối disulfua này có bản sao trong A (được đưa ra trong Skehel and Wiley, 2000. Ann Rev Biochem. 69:531-569; các ví dụ về các cấu trúc minh họa các cầu nối disulfua giữa các phân tử và nội phân tử và các axit amino

được bảo tồn khác và các vị trí tương ứng của chúng được mô tả trong, ví dụ, Gamblin et al 2004 (Science 303:1838-1842); cả hai được kết hợp ở đây để tham khảo. Với một người có trình độ trung bình trong lĩnh vực kỹ thuật này sẽ nhận ra rằng, rất cần phải bảo đảm sự sắp xếp tương tự các cầu nối disulfua thu được khi bào chế các HA thể khảm.

Peptit tín hiệu có thể là nguyên bản với chất ngưng kết hồng cầu, hoặc peptit tín hiệu có thể khác loại đối với trình tự ban đầu của chất ngưng kết hồng cầu được biểu hiện. HA thể khảm có thể bao gồm peptit tín hiệu từ loại, typ phụ cúm thứ nhất nhờ cân xứng với HA từ một hoặc nhiều hơn một loại, typ phụ, hoặc chủng cúm khác nhau. Ví dụ SP gốc thuộc typ phụ B của HA H1, H2, H3, H5, H6, H7, H9 hoặc cúm typ B có thể được sử dụng để biểu hiện HA trong hệ thực vật. Trong một vài phương án thực hiện theo sáng chế, SP có thể là cúm typ B, H1, H3 hoặc H5; hoặc thuộc typ phụ H1/Bri, H1/NC, H5/Indo, H3/Bri hoặc B/Flo.

SP cũng có thể không phải là gốc, ví dụ, từ protein có cấu trúc hoặc chất ngưng kết hồng cầu của virus khác với bệnh cúm, hoặc từ polypeptit của vi khuẩn, động vật hoặc thực vật. Ví dụ không giới hạn về peptit tín hiệu có thể được sử dụng là peptit của protein disulfua isomeraza ở cỏ linh lăng (PDI SP; các nucleotit 32-103 có số gia nhập Z11499; SEQ ID NO: 34; Fig.17) có trình tự axit amin:

MAKNVAIFGLLFSLLLLVPSQIFAEE (nucleotit 32-103; SEQ ID NO: 34)

Do đó sáng chế cũng đề xuất chất ngưng kết hồng cầu gây bệnh cúm thể khảm bao gồm peptit tín hiệu không nguyên bản, hoặc nguyên bản, và các axit nucleic ghi hóa các chất ngưng kết hồng cầu thể khảm này.

Sự gấp nếp chính xác các chất ngưng kết hồng cầu có thể có tầm quan trọng với sự ổn định protein, sự tạo thành các đa phân, sự tạo thành các VLP và chức năng của HA (khả năng để ngưng kết hồng cầu), trong số các đặc tính khác của chất ngưng kết hồng cầu cúm. Sự gấp nếp protein có thể bị ảnh hưởng bởi một hoặc nhiều yếu tố, gồm, nhưng không bị giới hạn, với trình tự protein, sự dồi dào protein tương đối, mức độ dồn đông trong tế bào, khả năng của các đồng yếu tố có thể liên kết hoặc tập hợp tạm thời với protein được gấp nếp, được gấp nếp một phần hoặc không được gấp nếp, khi có mặt một hoặc nhiều protein chaperon, hoặc tương tự.

Các protein sốc nhiệt (Hsp) hoặc các protein gây căng thẳng là các ví dụ về các protein chaperon, chúng có thể tham gia vào các quy trình tế bào khác nhau gồm sự tổng

hợp protein, vận chuyển nội bào, ngăn chặn sự gấp nếp không đúng, ngăn chặn sự tích tụ protein, gắn và tháo các phức protein, gấp nếp protein, và phân rã protein. Các ví dụ về các protein chaperon gồm, nhưng không giới hạn với, Hsp60, Hsp65, Hsp 70, Hsp90, Hsp100, Hsp20-30, Hsp10, Hsp100-200, Hsp100, Hsp90, Lon, TF55, các FKBP, xyclophilin, ClpP, GrpE, ubiquitin, calnexin, và các protein disulfua isomerase (tham khảo, 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); U.S. Patent No. 5,232,833). Như được mô tả ở đây, các protein chaperon, ví dụ, nhưng không chỉ giới hạn bởi, Hsp40 và Hsp70 có thể được sử dụng để bảo đảm sự gấp nếp của HA thể khảm.

Các ví dụ về Hsp70 gồm Hsp72 và Hsc73 từ các tế bào động vật có vú, DnaK từ vi khuẩn, cụ thể là vi khuẩn hiếu khí gram dương hình que như *Mycobacterium leprae*, *Mycobacterium tuberculosis*, và *Mycobacterium bovis* (như Bacille-Calmette Guerin: đề cập đến ở đây dưới dạng Hsp71). DnaK từ *Escherichia coli*, nấm men và các sinh vật chưa có nhân điển hình, và BiP và Grp78 từ các sinh vật có nhân thật, như *A. thaliana* (Lin et al. 2001 (Cell Stress and Chaperones 6:201-208). Ví dụ cụ thể về Hsp70 là Hsp70 là *A. thaliana* Hsp70 (được mã hóa bởi Ngân hàng gen số tham chiếu: AY120747.1). Hsp70 có khả năng liên kết đặc hiệu ATP cũng như các polypeptit và peptit không được gấp nếp, nhờ đó tham gia vào sự gấp nếp và không gấp nếp protein cũng như gắn và tháo các phức protein.

Các ví dụ về Hsp40 gồm DnaJ từ các sinh vật chưa có nhân điển hình như *E. coli* và vi khuẩn hiếu khí gram dương hình que và HsJ1, HDJ1 và Hsp40 từ các sinh vật có nhân điển hình, như cỏ linh lăng (Frugis et al., 1999. Plant Molecular Biology 40:397-408). Ví dụ cụ thể về Hsp40 là *M. sativa* MsJ1 (AJ000995.1 hoặc SEQ ID NO: 76). Hsp40 đóng vai trò làm chaperon phân tử trong sự gấp nếp protein, bền nhiệt và tái bản ADN, trong số các hoạt động khác của tế bào. FIG.41 thể hiện trình tự axit nucleic mã hóa Msj1 (SEQ ID NO: 76).

Trong số Hsps, Hsp70 và co-chaperon của nó, Hsp40, có liên quan đến sự ổn định sự dịch mã và các polypeptit được tổng hợp mới trước khi hoàn thành sự tổng hợp. Không mong muốn bị ràng buộc bởi lý thuyết, Hsp40 gắn với các mảng kỵ nước của các polypeptit không được gấp nếp (mới sinh hoặc mới được chuyển), do đó tạo thuận lợi cho sự tương tác của phức Hsp70-ATP với polypeptit. Sự thủy phân ATP dẫn tới sự tạo thành phức ổn định giữa polypeptit, Hsp70 và ADP, và giải phóng Hsp40. Sự kết hợp của phức

Hsp70-ADP với các mảng kỵ nước của polypeptit ngăn cản sự tương tác của chúng với các mảng kỵ nước khác, ngăn cản sự gấp nếp không chính xác và tạo thành các khối tích tụ với các protein khác (được đưa ra trong Hartl, FU. 1996. Nature 381:571-579).

Các protein chaperon gốc có thể có khả năng tạo thuận lợi cho việc gấp nếp chính xác các mức thấp của protein tái tổ hợp, nhưng dưới dạng các mức hiệu hiện tăng, sự dồi dào của các chaperon gốc có thể trở thành yếu tố hạn chế. Các mức biểu hiện cao của chất ngưng kết hồng cầu trong các lá được thâm nhiễm Agro có thể dẫn tới sự tích tụ các chất ngưng kết hồng cầu polypeptit trong dung dịch bào tương, và sự biểu hiện đồng thời của một hoặc nhiều hơn một protein chaperin như Hsp70, Hsp40 hoặc cả Hsp70 và Hsp40 có thể làm giảm mức chất ngưng kết hồng cầu polypeptit được tích tụ hoặc gấp nếp không chính xác, và làm tăng số lượng các polypeptit thể hiện các đặc tính của cấu trúc bậc ba và bậc bốn cho phép ngưng kết hồng cầu và/hoặc tạo thành các hạt tương tự virus. SEQ ID NO: 77 là trình tự axit nucleic của một phần của sản phẩm thiết kế số R850, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu phiên mã) tới EcoRI (ngay cùng chiều với nhân tố kết thúc NOS), mã hóa HSP40 (được gạch dưới). SEQ ID NO: 78 là trình tự axit nucleic của một phần sản phẩm thiết kế số R860, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu phiên mã) tới EcoRI (ngay cùng chiều với nhân tố kết thúc NOS), mã hóa HSP70 (được gạch dưới). SEQ ID NO: 79 là trình tự axit nucleic của một phần sản phẩm thiết kế số R870, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu phiên mã) tới EcoRI (ngay cùng chiều với nhân tố kết thúc NOS) mã hóa HSP40 (được gạch đậm) và HSP70 (được gạch dưới).

Do đó, sáng chế cũng đề xuất phương pháp sản xuất các VLP cúm thể khả ở thực vật, trong đó axit nucleic thứ nhất ghi hóa HA cúm thể khả được biểu hiện đồng thời với axit nucleic thứ hai mã hóa cheperon. Các axit nucleic thứ nhất và thứ hai có thể được đưa vào thực vật trong cùng một bước, hoặc có thể được đưa vào thực vật một cách tuần tự.

Có thể đánh giá cấu trúc và kích thước của các VLP bằng, ví dụ, thử nghiệm ngưng kết hồng cầu, kính hiển vi điện tử, hoặc bằng phương pháp sắc ký loại trừ kích thước.

Đối với phương pháp sắc ký loại trừ kích thước, toàn bộ các protein hòa tan có thể được chiết từ mô thực vật bằng cách làm đồng nhất mẫu (Polytron) vật liệu thực vật đã được nghiền-đông lạnh trong đệm chiết, và vật liệu không hòa tan được loại bỏ bằng cách ly tâm. Kết tủa bằng PEG cũng rất hữu ích. Protein hòa tan được định lượng, và dịch chiết

được cho đi qua cột Sephacryl™. Có thể sử dụng phẩm Blue Dextran 2000 làm tiêu chuẩn định cỡ. Sau khi sắc ký, các đoạn có thể được phân tích thêm nhờ thẩm miễn dịch để xác định đoạn protein bổ sung.

Sáng chế cũng đề xuất cây bao gồm axit nucleic mã hóa một, hoặc nhiều hơn một chất ngưng kết hồng cầu cúm thể khảm và axit nucleic mã hóa một hoặc nhiều hơn một cheperon.

Sáng chế cũng gồm các trình tự nucleotit:

SEQ ID NO: 63 (sản phẩm thiết kế 690; catxet biểu hiện H5/H1 thể khảm bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cở linh lăng và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, plastoxyanin ở cở linh lăng 3' UTR và trình tự kết thúc) và đoạn gạch dưới của SEQ ID NO:63 mã hóa SP, F'1, E1 trong H5/Indo-RB của H1/Bri-E2, F'2, F, TMD/CT của H5/Indo;

SEQ ID NO: 64 (sản phẩm thiết kế 691; catxet biểu hiện H5/H1 thể khảm bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cở linh lăng và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, plastoxyanin ở cở linh lăng 3' UTR và trình tự kết thúc) và đoạn gạch dưới của SEQ ID NO:64, mã hóa SP, F'1, E1, trong H5/Indo-E1, RB.E2 của H1/Bri-F'2, F, TMD/CT của H5/Indo;

SEQ ID NO: 65 (sản phẩm thiết kế 696; catxet biểu hiện H5/H1 thể khảm bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cở linh lăng và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, plastoxyanin ở cở linh lăng 3' UTR và trình tự kết thúc) và đoạn gạch dưới của SEQ ID NO:65 mã hóa PDI SP-F'1, E1 trong H1/NC-RB của H5/Indo-E2, F'2, F, TMD/CT của H1/NC;

SEQ ID NO: 68 (sản phẩm thiết kế 733; catxet biểu hiện SpPDI H1/Bri bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa peptit tín hiệu từ PDI, trình tự mã hóa chất ngưng kết hồng cầu có dạng H1 A/Brisbane/59/07 (H1N1), CPMV-HT 3' UTR và trình tự kết thúc NOS), và đoạn gạch dưới có SEQ ID NO:68, mã hóa PDI SP-F'1, E1, RB,E2, F'2, F, TMD/CT của H1/BRI;

SEQ ID NO: 69 (sản phẩm thiết kế 734; catxet biểu hiện H5/H1 thể khảm bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa chất ngưng

kết hồng cầu thể khảm, CPMV-HT 3' UTR và trình tự kết thúc NOS). Trình tự mã hóa HA thể khảm được gạch dưới, mã hóa HA thể khảm tương tự như SEQ ID NO:63;

SEQ ID NO: 71 (sản phẩm thiết kế số 736; catxet biểu hiện HA bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa peptit tín hiệu từ PDI, trình tự mã hóa chất ngưng kết hồng cầu có dạng H3 A/Brisbane/10/07 (H2N3), CPMV-HT 3' UTR và trình tự kết thúc NOS), và đoạn gạch dưới có SEQ ID NO: 71 mã hóa PDI SP-F'1, E1, RB,E2, F'2, F, TMD/CT của H3/Bri;

SEQ ID NO: 72 (sản phẩm thiết kế 737; catxet biểu hiện H5/H1 thể khảm bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, CPMV-HT 3' UTR và trình tự kết thúc NOS), và đoạn gạch dưới có SEQ ID NO:72 mã hóa PDI SP-F'1, E1, RB,E2, F'2, F, TMD/CT của H5/Indo;

SEQ ID NO: 74 (sản phẩm thiết kế 739; catxet biểu hiện HA bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa peptit tín hiệu từ PDI, trình tự mã hóa chất ngưng kết hồng cầu có dạng HA B/Florida/4/06, CPMV-HT 3' UTR và trình tự kết thúc NOS), và đoạn gạch dưới có SEQ ID NO:74, mã hóa PDI SP-F'1, E1, RB,E2, F'2, F, TMD/CT của B/Flo;

SEQ ID NO: 75 (sản phẩm thiết kế 734; catxet biểu hiện H5/B thể khảm bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu thể khảm, CPMV-HT 3' UTR và trình tự kết thúc NOS), và đoạn gạch dưới có SEQ ID NO:75 mã hóa PDI SP-F'1, E1, RB,E2, F'2, F của B/Flo-TND/CT của H5/Indo.

Sáng chế cũng gồm trình tự nucleotit lai giống dưới các điều kiện lai giống nghiêm ngặt với các đoạn được gạch dưới của một trong các SEQ ID NO:63-65, 68, 69, và 71-75 bất kỳ. Sáng chế cũng gồm trình tự nucleotit lai giống dưới các điều kiện lai giống nghiêm ngặt với sự bổ sung của các đoạn được gạch dưới của một trong các SEQ ID NO:63-65, 68, 69, và 71-75 bất kỳ. Các trình tự nucleotit này lai với các đoạn được gạch dưới có các SEQ ID NO:63-65, 68, 69, và 71-75, hoặc sự bổ sung các đoạn được gạch dưới có SEQ ID NO:63-65, 68, 69, và 71-75, mã hóa protein chất ngưng kết hồng cầu thể khảm, khi các dạng được biểu hiện VLP thể khảm, và VLP thể khảm gây ra sự sản xuất kháng thể khi được sử dụng cho đối tượng. Ví dụ, sự biểu hiện trình tự nucleotit trong tế bào thực vật tạo thành VLP thể khảm, và VLP thể khảm có thể được sử dụng để sản xuất kháng thể

có khả năng gắn HA, gồm HA, HA0, HA1, hoặc HA2 hoàn thiện của một hoặc nhiều loại hoặc typ phụ cúm. VLP thể khảm, khi được sử dụng cho đối tượng, gây ra sự đáp ứng miễn dịch.

Lai dưới các điều kiện lai khắc nghiệt đã được biết đến trong lĩnh vực kỹ thuật (tham khảo ví dụ Current Protocols in Molecular Biology, Ausubel et al, eds. 1995 and supplements; Maniatis và cộng sự, in Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory, 1982; Sambrook and Russell, in Molecular Cloning: A Laboratory Manual, xuất bản lần thứ 3 năm 2001; mỗi tài liệu được kết hợp ở đây để tham khảo). Một ví dụ về điều kiện lai nghiêm ngặt trên có thể nằm trong khoảng từ 16 đến 20 giờ lai trong 4 X SSC ở nhiệt độ 65°C, sau đó bằng cách rửa trong 0,1 X SSC ở nhiệt độ 65°C trong khoảng thời gian một giờ, hoặc 2 lần rửa trong 0,1 X SSC ở nhiệt độ 65°C mỗi lần trong khoảng thời gian 20 hoặc 30 phút. Mặt khác, điều kiện lai khắc nghiệt được dẫn chứng có thể qua đêm (trong thời gian từ 16 tới 20 giờ) trong formamit 50%, 4 X SSC ở nhiệt độ 42°C, sau đó rửa trong 0,1 X SSC ở nhiệt độ 65°C trong khoảng thời gian 1 giờ, hoặc rửa hai lần trong 0,1 X SSC ở nhiệt độ 65°C, mỗi lần rửa trong 30 hoặc 20 phút, hoặc qua đêm (trong thời gian từ 16 tới 20 giờ), hoặc lai trong đệm phosphat có nước Church (SDS 7%; đệm NaPO₄ 0,5M pH=7,2; EDTA 10mM) ở nhiệt độ 65°C, nhờ rửa 2 lần hoặc ở nhiệt độ 50°C trong 0,1 X SSC, SDS 0,1% mỗi lần rửa trong 20 hoặc 30 phút, hoặc rửa hai lần ở nhiệt độ 65°C trong 2 X SSC, SDS 0,1% mỗi lần rửa trong thời gian mỗi 20 hoặc 30 phút.

Ngoài ra, sáng chế gồm các trình tự nucleotit được đặc trưng có khoảng 70, 75, 80, 85, 87, 90, 91, 92, 93 94, 95, 96, 97, 98, 99, 100% hoặc bất cứ lượng nào trong khoảng đó, có tính đồng nhất về trình tự, hoặc tương đồng về trình tự, với trình tự nucleotit mã hóa HA thể khảm theo các đoạn gạch dưới của một trong số các trình tự nêu trong SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 68, SEQ ID NO: 69, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 74, SEQ ID NO: 75, trong đó trình tự nucleotit mã hóa protein chất ngưng kết hồng cầu mà khi được biểu hiện tạo thành VLP thể khảm, và VLP thể khảm dùng để sản xuất kháng thể. Ví dụ, sự biểu hiện trình tự nucleotit trong tế bào thực vật tạo thành VLP thể khảm, và VLP thể khảm có thể được sử dụng để sản xuất kháng thể có khả năng gắn HA, gồm HA, HA0, HA1, hoặc HA2 hoàn thiện. VLP, khi được sử dụng cho đối tượng, gây ra đáp ứng miễn dịch.

“Đáp ứng miễn dịch” thường đề cập đến sự đáp ứng với hệ thống miễn dịch thích ứng. Hệ thống miễn dịch thích ứng thường bao gồm đáp ứng thể dịch, và đáp ứng trung gian tế bào. Đáp ứng thể dịch là một khía cạnh về tính miễn dịch được trung gian bởi các kháng thể được tiết ra, được sản xuất trong các tế bào thuộc dòng tế bào lympho B (tế bào B). Các kháng thể được tiết ra liên kết với các kháng nguyên trên bề mặt của các vi khuẩn xâm chiếm (như virus hoặc vi khuẩn), báo hiệu cho chúng về sự phá hủy. Sự miễn dịch thể dịch thường được sử dụng để cập đến việc sản xuất kháng thể và các quy trình đi kèm với nó, cũng như các chức năng cảm ứng của kháng thể, gồm sự hoạt hóa tế bào Th2 và sản xuất xytokin, sự tạo ra tế bào ghi nhớ, sự thúc đẩy opsonin của thực bào, sự loại trừ mầm bệnh và các quy trình tương tự. Thuật ngữ “điều biến” hoặc “sự điều biến” hoặc đề cập tương tự đến sự tăng hoặc giảm một đáp ứng hoặc thông số cụ thể, như được xác định bởi các thử nghiệm khác nhau thường được biết đến, một vài thử nghiệm được dẫn chứng trong bản mô tả.

Đáp ứng trung gian tế bào là đáp ứng miễn dịch không gồm các kháng thể nhưng đúng hơn là gồm sự hoạt hóa các đại thực bào, tế bào giết tự nhiên (NK), các tế bào lympho T gây độc tế bào đặc hiệu kháng nguyên, và sự giải phóng các xytokin khác nhau đáp ứng với kháng nguyên. Miễn dịch trung gian tế bào thường được sử dụng để cập đến sự hoạt hóa tế bào Th, sự hoạt hóa tế bào Tc và các đáp ứng trung gian tế bào T. Đáp ứng trung gian tế bào có tầm quan trọng đặc biệt đáp ứng với sự nhiễm virus.

Ví dụ, sự gây ra các tế bào lympho T dương tính CD8 đặc hiệu kháng nguyên có thể được đo nhờ sử dụng thử nghiệm ELISPOT; sự kích thích các tế bào lympho T dương tính CD4 có thể được đo nhờ sử dụng thử nghiệm tăng sinh. Các độ chuẩn kháng thể chống bệnh cúm có thể được định lượng nhờ sử dụng thử nghiệm ELISA; các isotype đặc hiệu kháng nguyên hoặc các kháng thể phản ứng chéo cũng có thể được đo nhờ sử dụng các kháng thể kháng isotype (ví dụ kháng -IgG, IgA, IgE hoặc IgM). Các phương pháp và kỹ thuật để tiến hành các thử nghiệm này đã được biết đến trong lĩnh vực kỹ thuật.

Thử nghiệm ức chế sự ngưng kết hồng cầu (HI, hoặc HAI) cũng có thể được sử dụng để chứng minh hiệu quả của các kháng thể được gây ra bởi vacxin, hoặc chế phẩm vacxin bao gồm HA thể khảm hoặc VLP thể khảm có thể ngăn chặn sự dính kết của các tế bào máu đỏ (RBC) bởi HA tái tổ hợp. Các độ chuẩn kháng thể ức chế sự ngưng kết hồng cầu của các mẫu huyết thanh có thể được đánh giá bằng đĩa vi chuẩn độ HAI (Aymard và cộng sự 1973). Các hồng cầu thuộc các loại khác nhau có thể được sử dụng – ví dụ, ngựa,

gà tây, gà hoặc tương tự. Thử nghiệm này đưa ra thông tin gián tiếp về sự gắn trime HA lên bề mặt VLP, xác nhận sự thể hiện thích hợp các vị trí kháng nguyên trên các HA.

Các độ chuẩn HAI phản ứng chéo cũng có thể được sử dụng để chứng minh hiệu quả của đáp ứng miễn dịch đối với các chủng virus khác có liên quan với typ phụ vaccin. Ví dụ, huyết thanh từ đối tượng được miễn dịch bằng chế phẩm vaccin bao gồm HDC thuộc typ hoặc typ phụ cúm thứ nhất có thể được sử dụng trong thử nghiệm HAI với chủng thứ hai của virus nguyên vẹn hoặc các hạt tương tự virus, và độ chuẩn HAI được xác định.

Không mong muốn bị giới hạn bởi lý thuyết, khả năng HA liên kết với RBC từ các động vật khác nhau được thực hiện bởi ái lực của HA đối với axit sialic được liên kết bằng các liên kết $\alpha 2,3$ hoặc $\alpha 2,6$ và sự có mặt của các axit sialic này trên bề mặt của RBC. HA từ virus gây cúm ở ngựa và chim dính kết các hồng cầu từ tất cả các loài khác nhau, gồm, gà tây, gà, vịt, chuột lang, người, cừu, ngựa và bò cái; trong khi các HA ở người sẽ gắn với hồng cầu của gà tây, gà, vịt, chuột lang, người và cừu (Ito T. et al, 1997, *Virology*, 227:493-499; Medeiros R et al, 2001. *Virology* 289:74-85).

Sự có mặt của xytokin hoặc các mức xytokin có thể được định lượng. Ví dụ sự đáp ứng tế bào hỗ trợ T (Th1/Th2) sẽ được đặc trưng bởi phép đo các tế bào tiết ra IFN- γ và IL-4 sử dụng ELISA (ví dụ các kit thử BD Biosciences OptEIA). Các tế bào đơn nhân của máu ngoại vi (PBMC) hoặc các tế bào lách thu được từ đối tượng có thể được nuôi cấy, và phần nổi trên bề mặt được phân tích. Các tế bào lympho T cũng có thể được định lượng bằng cách phân loại tế bào được hoạt hóa bởi sự phát huỳnh quang (FACS), sử dụng các nhãn huỳnh quang đặc hiệu chất đánh dấu và các phương pháp đã được biết trong lĩnh vực kỹ thuật.

Thử nghiệm trung hòa vi lượng cũng có thể được thực hiện để mô tả đáp ứng miễn dịch trong đối tượng, tham khảo ví dụ các phương pháp của Rowe et al., 1973. Độ chuẩn trung hòa virus có thể thu được theo cách khác nhau, gồm: 1) đếm mảng phân giải (thử nghiệm mảng) sau khi cố định/nhuộm màu tinh thể tím cho các tế bào; 2) quan sát bằng kính hiển vi sự phân giải tế bào trong canh trường; 3) ELISA và dò bằng quang phổ kế protein virus NP (tương quan với sự lây nhiễm virus của các tế bào chủ).

Tính tương đồng hoặc tính đồng nhất trình tự có thể được xác định nhờ sử dụng chương trình so sánh trình tự, như được đề xuất trong DNASIS (ví dụ, sử dụng, nhưng không giới hạn, với các thông số sau đây: GAP penalty 5, # của đường chéo đỉnh 5, GAP

penalty 10 được cố định, k-tuple 2, khoảng trống nội 10, và kích thước cửa sổ 5). Tuy nhiên, các phương pháp sắp xếp trình tự khác để so sánh đã được biết trong lĩnh vực kỹ thuật ví dụ các 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 các đoạn bổ sung được lưu trữ trong máy tính của các thuật toán này (ví dụ GAP, BESTFIT, FASTA, và BLAST (Altschul et al., 1990, *J. Mol Biol* 215:403-410), hoặc bằng sự sắp xếp thủ công và kiểm tra bằng mắt thường. Trình tự axit amin hoặc axit nucleic có thể so sánh hoặc sắp thẳng hàng và trình tự tương đồng có thể được xác định bằng cách sử dụng bất cứ gói phần mềm nào đã biết trong lĩnh vực kỹ thuật, ví dụ MULTALIN (Corpet F., 1988, *Nucl. Acids Res.*, 16 (22), 10881-10890), BLAST, CLUSTAL hoặc tương tự, các trình tự lần lượt có thể được sắp thẳng hàng thủ công và các đặc điểm tương tự và khác biệt giữa các trình tự được xác định.

Một đoạn hoặc một phần protein, protein dung hợp hoặc polypeptit gồm peptit hoặc polypeptit bao gồm tập hợp con của phần phụ axit amino của một protein hoặc polypeptit cụ thể, được đề xuất có thể tạo ra VLP khi được biểu hiện. Đoạn đó có thể, ví dụ, bao gồm vùng kháng nguyên, vùng gây đáp ứng căng thẳng, hoặc vùng bao gồm vùng chức năng của protein hoặc polypeptit. Đoạn cũng có thể bao gồm vùng hoặc miền phổ biến với các protein thuộc họ chung, hoặc đoạn đó có thể gồm các trình tự axit amin đủ để nhận biết cụ thể một protein có độ dài đầy đủ từ đoạn mà nó được nhận.

Ví dụ, đoạn hoặc phần có thể bao gồm từ khoảng 60% tới khoảng 100%, độ dài của một protein có độ dài đầy đủ, hoặc giá trị bất kỳ nằm trong khoảng này, miễn là đoạn đó có thể tạo ra VLP thể khảm khi được biểu hiện. Ví dụ, từ khoảng 60% tới khoảng 100%, từ khoảng 70% tới khoảng 100%, từ khoảng 80% tới khoảng 100%, từ khoảng 90% tới khoảng 100%, từ khoảng 95% tới khoảng 100%, độ dài của một protein có độ dài đầy đủ, hoặc bất cứ giá trị nào giữa các khoảng đó. Mặt khác, đoạn hoặc phần đó có thể nằm trong khoảng từ 150 tới khoảng 500 axit amin, hoặc giá trị bất kỳ nằm trong khoảng này, phụ thuộc vào HA thể khảm, và miễn là đoạn đó có thể tạo ra VLP thể khảm khi được biểu hiện. Ví dụ, đoạn có thể nằm trong khoảng từ 150 tới khoảng 500 axit amin, hoặc giá trị bất kỳ nằm trong khoảng này, từ khoảng 200 tới khoảng 500 axit amin, hoặc giá trị bất kỳ nằm trong khoảng này, từ khoảng 250 tới khoảng 500 axit amin, hoặc giá trị bất kỳ nằm trong khoảng này, từ khoảng 300 tới khoảng 500 hoặc giá trị bất kỳ nằm trong khoảng này, từ khoảng 350 tới khoảng 500 axit amin, hoặc giá trị bất kỳ nằm trong khoảng này, từ

khoảng 400 tới khoảng 500 hoặc giá trị bất kỳ nằm trong khoảng này, từ khoảng 450 tới khoảng 500 hoặc giá trị bất kỳ nằm trong khoảng này, phụ thuộc vào HA thể khảm, và miễn là đoạn đó có thể tạo ra VLP thể khảm khi được biểu hiện. Ví dụ, khoảng 5, 10, 20, 30, 40 hoặc 50 axit amin, hoặc giá trị bất kỳ nằm trong khoảng này có thể được loại bỏ khỏi đầu tận cùng C, đầu tận cùng N hoặc cả đầu tận cùng N và C của protein HA thể khảm, miễn là đoạn có thể tạo ra VLP thể khảm khi được biểu hiện.

Sự đánh số axit amin trong bất cứ trình nào đã đưa ra có liên quan tới một trình tự cụ thể, tuy nhiên người có trình độ trung bình có thể dễ dàng xác định “tính tương đương” của một axit amino cụ thể trong trình tự dựa trên cấu trúc và/hoặc trình tự. Ví dụ, 6 axit amin đầu tận cùng N được loại bỏ, điều này sẽ làm thay đổi sự nhận dạng bằng số đặc trưng của axit amin (ví dụ, tương ứng với độ dài đầy đủ của protein), nhưng không làm thay đổi vị trí tương ứng của axit amin trong cấu trúc.

Sáng chế mô tả, nhưng không giới hạn với, sự biểu hiện axit nucleic mã hóa HA thể khảm trong bộ phận của cây, hoặc tế bào cây, và sự sản xuất các VLP bệnh cúm thể khảm, thích hợp để sản xuất vacxin. Các ví dụ về axit nucleic gồm, ví dụ, nhưng không giới hạn với, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 68, SEQ ID NO: 69, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 74, SEQ ID NO: 75.

Sáng chế cũng đề xuất thêm sự biểu hiện axit nucleic mã hóa HA thể khảm, ví dụ nhưng không giới hạn với SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 68, SEQ ID NO: 69, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 74, SEQ ID NO: 75 trong cây, bộ phận của cây, hoặc tế bào của cây, và sản xuất các thành phần vắc xin hoặc chất phản ứng gồm có các protein cấu trúc cúm tái tổ hợp tự gắn vào các cấu trúc protein đại phân tử cùng kiểu tạo miễn dịch và có chức năng, gồm các hạt cúm virus dưới đơn vị và VLP cúm thể khảm, trong các tế bào thực vật được biến nạp.

Do đó, sáng chế đề xuất các VLP thể khảm, và phương pháp sản xuất các VLP thể khảm trong hệ biểu hiện thực vật, từ sự biểu hiện của protein thể khảm có vỏ bọc đơn.

Axit nucleic mã hóa HA thể khảm thuộc typ phụ cúm, ví dụ SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 68, SEQ ID NO: 69, SEQ ID NO: 71, SEQ ID NO: 72, SEQ ID NO: 74, SEQ ID NO: 75 có thể được tổng hợp bởi sự phiên mã ngược và phản ứng chuỗi polymeraza (PCR) sử dụng HA ARN. Ví dụ, ARN có thể được phân lập từ H1/NC, H1/Bri, H3/Bri, B/Flo hoặc H5/Indo, hoặc từ các tế bào bị nhiễm các typ hoặc

typ phụ virus cúm này hoặc virus cúm khác. Để phiên mã ngược và PCR, mỗi oligonucleotit đặc hiệu cho HA ARN có thể được sử dụng. Ngoài ra, axit nucleic mã hóa HA thể khảm có thể được tổng hợp bằng hóa học sử dụng các phương pháp như đã được biết bởi người có trình độ trung bình trong lĩnh vực kỹ thuật này.

Sáng chế còn đề cập đến sản phẩm thiết kế gen bao gồm axit nucleic mã hóa HA, như được mô tả ở trên, được liên kết có hiệu quả với yếu tố điều hòa có hoạt tính với thực vật. Các ví dụ của các yếu tố điều hòa có hoạt tính trong tế bào thực vật và có thể được sử dụng theo sáng chế gồm nhưng không giới hạn với vùng điều hòa plastocyanin (US 7125978; được kết hợp ở đây để tham khảo), hoặc vùng điều hòa Ribulose 1,5-bisphosphat carboxylase/oxygenase (RuBisCO; US 4962028; được kết hợp ở đây để tham khảo), protein gắn với chlorophyll a/b (CAB; Leutwiler et al; 1986; được kết hợp ở đây để tham khảo), ST-LS1 (gắn với phức giải phóng oxy của hệ thống quang hợp II và được mô tả bởi Stockhaus et al. 1987, 1989; được kết hợp ở đây để tham khảo).

Sản phẩm thiết kế gen theo sáng chế cũng có thể bao gồm trình tự khởi đầu phiên mã cơ định trực tiếp biểu hiện gen được liên kết có hiệu quả với trình tự khởi đầu phiên mã thông qua các phần khác nhau của cây và liên tục trong suốt quá trình phát triển của cây. Ví dụ không giới hạn với trình tự khởi đầu phiên mã cơ định được kết hợp với sản phẩm phiên mã CaMV 35S (ví dụ Odell et al., 1985, Nature, 313: 810-812, được kết hợp ở đây để tham khảo).

Ví dụ về trình tự bao gồm vùng điều hòa plastocyanin là trình tự 5' đến trình tự gạch dưới mã hóa peptit tín hiệu PDI có SEQ ID NO: 58. Yếu tố điều hòa hoặc vùng điều hòa có thể cải thiện sự dịch mã của trình tự nucleotit mà nó liên kết có hiệu quả, tại đó trình tự nucleotit có thể mã hóa protein hoặc polypeptit. Ví dụ khác của vùng điều hòa, được nhận từ các vùng không được dịch mã của virus khảm Cây đậu đũa (CPMV), nó có thể được sử dụng để dịch mã ưu tiên trình tự nucleotit mà nó liên kết có hiệu quả. Vùng điều hòa CPMV này được phát hiện trong hệ CPMV có thể dịch mã cao (CPMV-HT; tham khảo, ví dụ, Sainsbury et al, 2008, Plant Physiology 148: 1212-1218; Sainsbury et al., 2008 Plant Biotechnology Journal 6:82-92; cả hai được kết hợp ở đây để tham khảo).

Do đó, một khía cạnh của sáng chế đề xuất axit nucleic bao gồm vùng điều hòa được liên kết có hiệu quả với trình tự mã hóa HA cúm thể khảm. Vùng điều hòa có thể là nhân tố điều hòa plastocyanin, và HA cúm thể khảm có thể bao gồm các vùng phụ từ các typ,

typ phụ hoặc chủng cúm H5/Indo, H1/Bri, H3/Bri, H1/NC, B/Flo. Trình tự axit nucleic bao gồm nhân tố điều hòa plastoxyanin và HA cúm thể khả được minh họa ở đây bằng các SEQ ID NO: 63 và 64. Trình tự axit nucleic bao gồm nhân tố điều hòa 35S và HA cúm thể khả được minh họa ở đây bằng SEQ ID NO: 68, 69 và 71-75.

Theo một khía cạnh khác, sáng chế đề xuất axit nucleic bao gồm vùng điều hòa CPMV và HA cúm thể khả, bao gồm các vùng phụ từ các typ, typ phụ hoặc chủng cúm H5/Indo, H1/Bri, H3/Bri, H1/NC, B/Flo. Trình tự axit nucleic bao gồm nhân tố điều hòa CPMP và HA thể khả được minh họa ở đây bằng SEQ ID NO: 66-69 và 71-75.

Các VLP cúm thể khả được tạo ra ở thực vật sinh chồi từ màng tế bào và chế phẩm lỏng của các VLP thể khả mà màng tế bào hoặc loại mô thực vật từ khi chúng được tạo ra. Các VLP được tạo ra theo sáng chế bao gồm HA của một hoặc nhiều hơn hai typ hoặc typ phụ của bệnh cúm, được tạo phức với các lipit có nguồn gốc từ thực vật. Các lipit của cây có thể kích thích các tế bào miễn dịch đặc hiệu và làm tăng đáp ứng miễn dịch gây ra.

Lipit thực vật như PC (phosphatidyl cholin) và PE (phosphatidyl etanolamin), cũng như các glycosphingolipit có thể liên kết với các phân tử CD1 được biểu hiện bởi các tế bào miễn dịch của động vật có vú như các tế bào trình diện kháng nguyên (các APC) giống các tế bào tua và các đại thực bào và các tế bào khác gồm các lympho B và T trong tuyến ức và gan (được đưa ra trong Tsuji M., 2006 Cell Mol Life Sci 63:1889-98). Các phân tử CD1 có cấu trúc tương tự với các phân tử phức hợp tương thích mô chính (MHC) trong lớp I và vai trò của chúng để trình diện kháng nguyên glycolipit với các tế bào NKT (tế bào T giết tự nhiên). Dựa trên sự hoạt hóa, các tế bào NKT hoạt hóa các tế bào miễn dịch bẩm sinh như các tế bào NK và tế bào tua và cũng hoạt hóa các tế bào miễn dịch thích ứng giống các tế bào B và tế bào T sản xuất kháng thể.

Các phytosterol thể hiện trong VLP cúm được tạo phức với lớp kép lipit, như vỏ có nguồn gốc từ màng tế bào có thể tạo ra chế phẩm vaccin có lợi. Không muốn bị giới hạn bởi bất cứ lý thuyết nào, các VLP được tạo từ thực vật, gồm VLP bao gồm HA thể khả, được tạo phức với lớp kép lipit, như vỏ bọc có nguồn gốc từ màng tế bào, có thể gây ra phản ứng miễn dịch mạnh hơn so với các VLP được tạo ra trong các hệ thống biểu hiện khác, và có thể giống với phản ứng miễn dịch được gây ra bởi các vaccin virus sống hoặc đã được giảm độc lực.

Do đó, theo một số phương án thực hiện, sáng chế đề xuất VLP bao gồm HA thể khảm, được tạo phức với lớp kép lipit có nguồn gốc từ thực vật. Theo một vài phương án thực hiện lớp kép lipit có nguồn gốc từ thực vật bao gồm vỏ của VLP.

VLP được tạo ra trong thực vật có thể gồm HA thể khảm bao gồm các N-glycan đặc trưng ở thực vật. Do đó, sáng chế cũng đề xuất VLP bao gồm HA thể khảm có các N-glycan đặc trưng ở thực vật.

Hơn nữa, sự biến đổi N-glycan ở thực vật đã được biết đến (tham khảo ví dụ WO 2008/151440; được kết hợp ở đây để tham khảo) và HA thể khảm có các N-glycan được biến đổi có thể được sản xuất. HA thể khảm bao gồm mẫu glycosyl hóa được biến đổi, ví dụ với các N-glycan được fucosyl hóa giảm, xylosyl hóa, hoặc cả hai, được fucosyl hóa và xylosyl hóa có thể thu được, hoặc HA thể khảm có mẫu glycosyl hóa được biến đổi có thể thu được, trong đó protein thiếu fucosyl hóa, xylosyl hóa, hoặc cả hai, và bao gồm sự galatosyl hóa tăng. Ngoài ra, sự điều biến các biến đổi sau dịch mã, ví dụ, sự bổ sung galatoza kết thúc có thể dẫn tới làm giảm sự fucosyl hóa và xylosyl hóa HA thể khảm được biểu hiện khi so sánh với cây kiểu đại biểu hiện HA thể khảm.

Ví dụ, không xem xét đến sự hạn chế, sự tổng hợp HA thể khảm có mẫu glycosyl hóa được biến đổi có thể đạt được bằng cách biểu hiện đồng thời protein đang nói đến cùng với trình tự nucleotit mã hóa beta-1.4galactosyltransferaza (GalT), ví dụ, nhưng không hạn chế với GalT ở động vật có vú, hoặc GalT ở người tuy nhiên cũng có thể sử dụng GalT từ các nguồn khác. Vùng xúc tác của GalT cũng có thể dung hợp với vùng CTS (tức là, đuôi tế bào chất, vùng xuyên màng, vùng gốc) của N-acetylglucosaminyl transferaza (GNT1), để tạo ra enzym lai GNT1-GalT, và enzym lai có thể được biểu hiện đồng thời với HA. HA cũng có thể được biểu hiện đồng thời cùng với trình tự nucleotit mã hóa N-acetylglucosaminyltransferaza III (GnT-III), ví dụ nhưng không bị giới hạn với GnT-III ở động vật có vú hoặc GnT-III ở người, cũng có thể sử dụng GnT-III từ các nguồn khác. 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 gồm các VLP bao gồm HA thể khảm có các N-glycan được biến đổi.

Không mong muốn bị ràng buộc bởi lý thuyết, sự có mặt của các N-glycan thực vật trên HA thể khảm có thể kích thích đáp ứng miễn dịch bằng cách thúc đẩy sự liên kết của

HA bởi các tế bào trình diện kháng nguyên. Sự kích thích đáp ứng miễn dịch sử dụng N glycan thực vật được đề xuất bởi Saint-Jore-Dupas et al. (Trends Biotechnol 25: 317-23, 2007). Hơn nữa, sự thích ứng của VLP có thể tạo thuận lợi cho sự trình diện kháng nguyên, và làm tăng hiệu quả tá được của VLP khi được tạo phức với lớp lipit có nguồn gốc từ thực vật.

“Vùng điều hòa”, “nhân tố điều hòa” hoặc “trình tự khởi đầu phiên mã” được hiểu là một phần của axit nucleic thông thường, nhưng không thường xuyên, ngược chiều của vùng mã hóa protein của gen, có thể gồm có ADN hoặc ARN, cả ADN và ARN. Khi vùng điều hòa được hoạt hóa, và trong sự kết hợp có hiệu quả, hoặc được liên kết có hiệu quả, với gen được quan tâm, điều này có thể dẫn đến sự biểu hiện gen được quan tâm. Nhân tố điều hòa có thể có khả năng điều chỉnh tính đặc hiệu của cơ quan, hoặc kiểm soát sự hoạt hóa gen theo thời gian hoặc theo sự tiến triển. “Vùng điều hòa” gồm các nhân tố trình tự khởi đầu phiên mã, các nhân tố trình tự khởi đầu phiên mã trung tâm thể hiện hoạt tính trình tự khởi đầu phiên mã cơ bản, các nhân tố cảm ứng với sự đáp ứng các kích thích bên ngoài, các nhân tố điều chỉnh hoạt tính trình tự khởi đầu phiên mã như các nhân tố điều hòa âm tính hoặc các nhân tố tăng cường phiên mã. “Vùng điều hòa”, như được sử dụng ở đây, cũng gồm các nhân tố hoạt hóa sau khi phiên mã, ví dụ, nhân tố điều hòa điều chỉnh sự biểu hiện gen như các nhân tố tăng cường phiên mã và dịch mã, các chất ức chế sự phiên mã và dịch mã, các trình tự hoạt hóa ngược chiều, và các yếu tố quyết định tính không ổn định của mARN. Một vài nhân tố điều hòa này có thể được đặt ở đầu gần vùng mã hóa.

Trong ngữ cảnh của bản mô tả, thuật ngữ “nhân tố điều hòa” hoặc “vùng điều hòa” thường đề cập đến trình tự ADN, thường, nhưng không thường xuyên, ngược chiều (5') với trình tự mã hóa gen cấu trúc, kiểm soát sự biểu hiện của vùng mã hóa bằng cách tạo ra sự nhận biết ARN polymeraza và/hoặc các nhân tố khác cần cho sự phiên mã để bắt đầu tại ở một vị trí cụ thể. Tuy nhiên, điều này được hiểu rằng các trình tự nucleotit khác, được định vị trong intron, hoặc đầu 3' của trình tự có thể góp phần vào việc điều hòa sự biểu hiện vùng mã hóa được quan tâm. Một ví dụ về nhân tố điều hòa tạo ra sự nhận biết ARN polymeraza hoặc các yếu tố phiên mã khác để bảo đảm sự khởi đầu tại vị trí đặc biệt là nhân tố trình tự khởi đầu phiên mã. Hầu hết, nhưng không phải là tất cả, các nhân tố trình tự khởi đầu phiên mã của sinh vật nhân chuẩn chứa hộp TATA, trình tự axit nucleic được bảo tồn gồm các cặp bazơ adenosin và thymidin nucleotit thường đặt ở vị trí khoảng

25 cặp bazơ ngược chiều với vị trí khởi đầu phiên mã. Nhân tố trình tự khởi đầu phiên mã bao gồm nhân tố trình tự khởi đầu phiên mã cơ bản, chịu trách nhiệm khởi đầu sự phiên mã, cũng như các nhân tố điều hòa khác (như được liệt kê ở trên) biến đổi sự biểu hiện gen.

Có các loại vùng điều hòa khác nhau, gồm các vùng được điều hòa sự phát triển, cảm ứng được hoặc cơ bản. Vùng điều hòa được điều hòa sự phát triển, hoặc điều chỉnh sự biểu hiện khác nhau của gen bằng cách kiểm soát nó, được hoạt hóa trong các bộ phận hoặc mô nhất định của cơ quan tại thời gian đặc biệt trong sự phát triển của cơ quan hoặc mô. Tuy nhiên, một số vùng điều hòa được điều hòa theo sự tiến triển có thể ưu tiên có hoạt tính trong các cơ quan hoặc mô nhất định ở các giai đoạn phát triển đặc trưng, chúng cũng có thể có hoạt tính trong phương thức điều hòa sự phát triển, hoặc ở mức cơ bản trong các cơ quan hoặc mô bên trong thực vật. Các ví dụ về các vùng điều hòa đặc hiệu mô, ví dụ vùng điều hòa đặc hiệu, gồm trình tự khởi đầu phiên mã napin, và trình tự khởi đầu phiên mã cruciferin (Rask et al., 1998, *J. Plant Physiol.* 152: 595-599; Bilodeau và cộng sự, 1994, *Plant Cell* 14: 125-130). Ví dụ về trình tự khởi đầu phiên mã đặc hiệu lá gồm trình tự khởi đầu phiên mã plastoxyanin (tham khảo, ví dụ SEQ ID NO: 58); US 7125978, được kết hợp ở đây để tham khảo).

Vùng điều hòa gây cảm ứng là vùng có khả năng hoạt hóa gián tiếp hoặc trực tiếp sự phiên mã một hoặc nhiều trình tự ADN hoặc gen đáp ứng với tác nhân gây cảm ứng. Khi không có mặt tác nhân gây cảm ứng các trình tự ADN hoặc gen không được phiên mã. Thông thường yếu tố protein liên kết đặc hiệu với vùng điều hòa gây cảm ứng để hoạt hóa sự phiên mã có thể có mặt dưới dạng bất hoạt, sau đó được biến đổi trực tiếp hoặc gián tiếp thành dạng hoạt hóa bởi tác nhân gây cảm ứng. Tuy nhiên, yếu tố protein cũng có thể vắng mặt. Tác nhân gây cảm ứng có thể là chất hóa học như protein, chất chuyển hóa, chất điều hòa sinh trưởng, thuốc diệt cỏ hoặc hợp chất phenolic hoặc căng thẳng sinh lý học phụ thuộc trực tiếp vào các nhân tố nóng, lạnh, muối, hoặc độc tố hoặc gián tiếp thông qua sự hoạt hóa mầm bệnh hoặc tác nhân gây bệnh như virus. Tế bào cây chứa vùng điều hòa gây cảm ứng được phơi nhiễm với tác nhân gây cảm ứng bằng cách đưa tác nhân gây cảm ứng vào tế bào hoặc cây như bằng cách phun, tưới nước, làm nóng hoặc các phương pháp tương tự. Các yếu tố điều hòa gây cảm ứng có thể có nguồn gốc từ gen thực vật hoặc gen không phải ở thực vật (ví dụ, Gatz, C. và Lenk, I.R.P., 1998, *Trends Plant Sci.* 3, 352-358; được kết hợp ở đây để tham khảo). Ví dụ, các trình tự khởi đầu phiên mã

cảm ứng có hiệu quả gồm, nhưng không giới hạn với, trình tự khởi đầu phiên mã cảm ứng tetracyclin (Gatz, C., 1997, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 48, 89-108; được kết hợp ở đây để tham khảo), trình tự khởi đầu phiên mã cảm ứng steroid (Aoyama, T. và Chua, N.H., 1997, *Plant J.* 2, 397-404; được kết hợp ở đây để tham khảo) và trình tự khởi đầu phiên mã cảm ứng etanol (Salter, M.G., và cộng sự, 1998, *Plant Journal* 16, 127-132; Caddick, M.X., et al, 1998, *Nature Biotech.* 16, 177-180, được kết hợp ở đây để tham khảo) các gen IB6 và CKI1 cảm ứng với xytokinin (Brandstatter, I. và Kieber, J.J., 1998, *Plant Cell* 10, 1009-1019; Kakimoto, T., 1996, *Science* 274, 982-985; được kết hợp ở đây để tham khảo) và yếu tố cảm ứng với auxin, DR5 (Ulmasov, T., et al., 1997, *Plant Cell* 9, 1963-1971; được kết hợp ở đây để tham khảo).

Vùng điều hòa cơ định hướng tới sự biểu hiện gen trong các phần khác nhau của cây và thông qua sự phát triển liên tục của cây. Các ví dụ về các nhân tố điều hòa cơ định đã biết gồm các trình tự khởi đầu phiên mã liên quan đến sản phẩm phiên mã CaMV 35S. (Odell et al., 1985, *Nature*, 313: 810-812), cây lúa actin 1 (Zhang et al, 1991, *Plant Cell*, 3: 1155-1165), actin 2 (An et al., 1996, *Plant J.*, 10: 107-121), hoặc tms 2 (U.S. 5428147, được kết hợp ở đây để tham khảo), và các gen triosephosphat isomerase 1 (Xu et al., 1994, *Plant Physiol.* 106: 459-467), gen ubiquitin 1 ở ngô (Cornejo et al, 1993, *Plant Mol. Biol.* 29: 637-646), các gen *Arabidopsis* ubiquitin 1 và 6 (Holtorf et al, 1995, *Plant Mol. Biol.* 29: 637-646), và gen 4A của yếu tố khởi đầu dịch mã ở cây thuốc lá (Mandel et al, 1995 *Plant Mol. Biol.* 29: 995-1004). Thuật ngữ "cơ định" được sử dụng trong bản mô tả không nhất thiết phải chỉ ra rằng gen nhờ sự kiểm soát vùng điều hòa cơ định được biểu hiện tại mức tương tự trong tất cả các loại tế bào, nhưng gen được biểu hiện trong nhiều loại tế bào khác nhau mặc dù thường quan sát thấy nhiều biến đổi. Các yếu tố điều hòa cơ định có thể được gắn với các trình tự khác để cải thiện hơn sự phiên mã và/hoặc dịch mã trình tự nucleotit mà chúng được liên kết có hiệu quả. Ví dụ, hệ thống CPMV-HT có nguồn gốc từ các vùng không được dịch mã của virus khảm đậu đũa (COMV) và chứng minh sự dịch mã được cải thiện của trình tự mã hóa được kết hợp.

“Gốc” nghĩa là trình tự axit nucleic hoặc axit amino ở dạng gốc, hoặc “thuần chủng”.

"Được liên kết có hiệu quả" nghĩa là các trình tự cụ thể, ví dụ nhân tố điều hòa và vùng mã hóa được quan tâm, tương tác trực tiếp hoặc gián tiếp để thực hiện chức năng được dự định, như sự trung gian hoặc điều hòa sự biểu hiện gen. Sự tương tác của các

trình tự được liên kết hiệu quả có thể, ví dụ, được trung gian bởi các protein tương tác với các trình tự được liên kết hiệu quả.

Một hoặc nhiều hơn một trình tự nucleotit theo sáng chế có thể được biểu hiện trong vật chủ cây thích hợp bất kỳ được biến nạp bằng trình tự nucleotit, hoặc sản phẩm thiết kế, hoặc các vectơ theo sáng chế. Các ví dụ về vật chủ thích hợp gồm, nhưng không giới hạn với, các cây trồng nông nghiệp gồm cỏ linh lăng, cải dầu canola, *Brassica* spp., ngô, *Nicotiana* spp., cỏ linh lăng, khoai tây, nhân sâm, đậu Hà Lan, yến mạch, lúa, đậu tương, lúa mì, lúa mạch, cây hướng dương, bông và các loài cây tương tự.

Một hoặc nhiều sản phẩm thiết kế di truyền thể khả năng theo sáng chế có thể bao gồm thêm vùng không được dịch mã 3'. Vùng không được dịch mã 3' đề cập tới một phần của gen bao gồm đoạn ADN chứa tín hiệu polyadenyl hóa và các tín hiệu điều hòa bất kỳ khác có khả năng tác động đến việc xử lý mARN hoặc biểu hiện gen. Tín hiệu polyadenyl hóa thường được mô tả bằng tác dụng bổ sung axit polyadenylic theo dấu đầu kết thúc 3' của tiền chất mARN. Các tín hiệu polyadenyl hóa thường được nhận biết nhờ sự có mặt tương đồng với dạng tiêu chuẩn 5' AATAAA-3' mặc dù những thay đổi này là không phổ biến.

Các ví dụ không giới hạn của các vùng 3' thích hợp là các vùng không dịch mã được phiên mã 3' chứa tín hiệu polyadenyl hóa của khối u *Agrobacterium* gây ra các gen plasmit (Ti), như gen nopalin synthaza (NOS), các gen thực vật như các gen protein dự trữ ở cây đậu tương, dưới đơn vị nhỏ của các gen ribuloza-1, 5-bisphosphat carboxylaza (ssRUBISCO; US 4962028; được kết hợp vào đây để tham khảo), trình tự khởi đầu phiên mã được sử dụng trong sự điều hòa biểu hiện plastoxyanin, được mô tả trong US 7125978 (được kết hợp ở đây để tham khảo).

Một hoặc nhiều sản phẩm thiết kế gen thể khả năng theo sáng chế cũng có thể gồm các yếu tố tăng cường khác, hoặc các yếu tố tăng cường sự dịch mã hoặc phiên mã, cũng được cần đến. Các yếu tố tăng cường có thể được định vị 5' hoặc 3' với trình tự được phiên mã. Những vùng yếu tố tăng cường này đã được biết bởi người có trình độ trung bình trong lĩnh vực kỹ thuật này, và có thể gồm codon khởi đầu ATG, các trình tự liền kề hoặc trình tự tương tự. Codon khởi đầu, nếu có mặt, có thể cùng pha với khung đọc ("trong khung") có trình tự mã hóa tạo ra sự dịch mã chính xác trình tự được phiên mã.

Để hỗ trợ việc nhận biết các tế bào cây được biến nạp, các sản phẩm thiết kế theo sáng chế có thể được thao tác gồm các chỉ thị chọn lọc của cây. Các chỉ thị chọn lọc hữu

ích gồm các enzym tạo ra tính kháng với các chất hóa học như thuốc kháng sinh ví dụ, gentamycin, hygromycin, kanamycin, hoặc thuốc diệt cỏ như phosphinothryxin, glyphosat, closulfuron, và các chất tương tự. Tương tự như vậy, tạo ra enzym để tạo ra hợp chất có thể nhận biết nhờ sự thay đổi màu sắc như GUS (betaglucuronidaza), hoặc sự phát quang, như luciferaza hoặc GFP, có thể được sử dụng.

Ngoài ra phần được xem xét theo sáng chế là các cây chuyển gen, các tế bào hoặc hạt cây chứa sản phẩm thiết kế gen thể thể khả năng theo sáng chế. Các phương pháp tạo ra toàn bộ cây từ các tế bào cây cũng đã được biết trong lĩnh vực kỹ thuật. Nói chung, các tế bào cây đã biến nạp được nuôi cấy trong môi trường thích hợp, có thể chứa các chất chọn lọc như thuốc kháng sinh, trong đó sử dụng các chỉ thị chọn lọc để nhận biết dễ dàng các tế bào cây đã biến nạp. Một lần nữa tạo thành mô sẹo, sự tạo thành chồi có thể được hỗ trợ nhờ sử dụng các hoocmôn thích hợp của cây phù hợp phù hợp với các phương pháp đã biết và các chồi được chuyển vào môi trường ra rễ để tái sinh cây. Sau đó các cây có thể được sử dụng để thiết lập các thể hệ lặp đi lặp lại, từ hạt hoặc sử dụng các kỹ thuật nhân giống sinh dưỡng. Các cây chuyển gen cũng có thể được tạo ra mà không sử dụng nuôi cấy mô.

Phần cũng được xem xét theo sáng chế là các cây chuyển gen, các cây, nấm men, vi khuẩn, nấm, côn trùng và các tế bào động vật chứa sản phẩm thiết kế gen thể thể khả năng bao gồm axit nucleic mã hóa HA hoặc HA0 thể thể khả năng, tái tổ hợp để sản xuất VLP, phù hợp với theo sáng chế.

Các yếu tố điều hòa theo sáng chế cũng có thể được kết hợp với vùng mã hóa được quan tâm để biểu hiện trong phạm vi các sinh vật chủ có thể tuân theo sự biến nạp, hoặc biểu hiện tạm thời. Các sinh vật này gồm, nhưng không giới hạn với cây, cả cây một lá mầm và cây hai lá mầm, ví dụ nhưng không giới hạn với ngô, cây ngũ cốc, lúa mì, lúa mạch, yến mạch, *Nicotiana* spp, *Brassica* spp, đậu tương, đậu, đậu Hà Lan, cỏ linh lăng, cà chua, khoai tây, nhân sâm và *Arabidopsis*.

Phương pháp để biến nạp ổn định, và tái sinh các sinh vật này được thiết lập trong lĩnh vực kỹ thuật và đã biết bởi người có trình độ trung bình trong lĩnh vực kỹ thuật này. Phương pháp thu được các cây biến nạp và được tái sinh không được quyết định theo sáng chế.

“Biến nạp” nghĩa là sự truyền thông tin di truyền giữa các loài khác nhau (trình tự nucleotit) được biểu thị dưới dạng kiểu gen, kiểu hình hoặc cả hai. Sự truyền thông tin di truyền giữa các loài khác nhau từ sản phẩm thiết kế thể khảm đến vật chủ có thể di truyền và truyền thông tin di truyền được xem xét ổn định, hoặc truyền thông tin di truyền tạm thời và truyền thông tin di truyền không thể thừa kế.

Thuật ngữ "vật chất của cây", có nghĩa là vật liệu bất kỳ có nguồn gốc từ thực vật. Vật chất của cây có thể bao gồm toàn bộ cây, mô, tế bào, hoặc phần bất kỳ của nó. Hơn nữa, vật chất của cây có thể bao gồm các thành phần nội bào, các thành phần ngoại bào của cây, dịch chiết lỏng hoặc rắn của cây, hoặc hỗn hợp của chúng. Hơn nữa, vật chất của cây có thể bao gồm các cây, tế bào cây, mô, dịch chiết lỏng, hoặc tổ hợp của chúng, từ các lá cây, thân, quả, rễ hoặc sự kết hợp của chúng. Vật chất của cây có thể bao gồm cây hoặc một phần của nó không phụ thuộc vào các bước xử lý bất kỳ. Bộ phận của cây có thể bao gồm vật chất của cây. Tuy nhiên, cũng dự tính được rằng vật liệu của cây tùy thuộc vào các bước xử lý tối thiểu như được định rõ dưới đây, hoặc xử lý nghiêm ngặt hơn, gồm sự tinh sạch protein một phần hoặc toàn bộ sử dụng các kỹ thuật thường được biết trong lĩnh vực kỹ thuật gồm, nhưng không giới hạn với phương pháp sắc ký, hiện tượng điện chuyển và các phương pháp tương tự.

Thuật ngữ "xử lý tối thiểu" nghĩa là vật chất của cây, ví dụ, cây hoặc một phần của nó bao gồm protein được quan tâm được tinh chế một phần để thu được dịch chiết của cây, chất đồng chất, một phần nhỏ chất đồng chất của cây hoặc chất tương tự (tức là được xử lý tối thiểu). Sự tinh chế một phần có thể bao gồm, nhưng không giới hạn với sự phá vỡ các cấu trúc tế bào của cây bằng cách đó tạo ra chế phẩm bao gồm các thành phần cây có thể hòa tan, và các phần của cây không hòa tan có thể được tách ra ví dụ, nhưng không giới hạn, bằng cách ly tâm, lọc hoặc kết hợp giữa chúng. Về mặt này, các protein được tiết ra trong không gian ngoại bào của lá hoặc các mô khác có thể dễ dàng thu được bằng cách sử dụng sự chiết ly tâm hoặc chân không, hoặc các mô có thể được chiết nhờ áp suất bằng cách chuyển qua các trục lăn hoặc nghiền hoặc các phương pháp tương tự để nén hoặc giải phóng protein không chứa trong không gian ngoại bào. Xử lý tối thiểu cũng có thể gồm sự điều chế dịch chiết thô của các protein hòa tan, do các chế phẩm này sẽ có sự nhiễm tạp không đáng kể từ các sản phẩm thứ cấp của cây. Hơn nữa, xử lý tối thiểu có thể gồm dịch chiết chứa nước của protein hòa tan từ lá, sau đó bằng cách kết tủa với muối

thích hợp bất kỳ. Các phương pháp khác có thể gồm sự thẩm ướt quy mô lớn và chiết nước ép để cho phép sử dụng trực tiếp dịch chiết.

Vật chất của cây, dưới dạng vật liệu cây hoặc mô có thể được phân phối bằng miệng đến đối tượng. Vật chất của cây có thể được sử dụng dưới dạng một phần bổ sung theo chế độ ăn, cùng với các thức ăn khác, hoặc được kết nang. Vật chất của cây hoặc mô cũng được tập trung để cải thiện hoặc làm tăng vị ngon, hoặc được tạo ra cùng với các vật liệu, thành phần, hoặc các tá dược khác, nếu cần.

Các ví dụ về đối tượng hoặc sinh vật đích có thể sử dụng các VLP theo sáng chế gồm, nhưng không chỉ giới hạn ở, con người, động vật linh trưởng, chim, chim nước, chim di cư, chim cú, vịt, ngan, ngỗng, gà vịt, gà, lợn, cừu, ngựa, ngựa, lạc đà, thú thuộc họ chó, chó, thú thuộc họ mèo, mèo, hổ, báo, cây hương, chồn vizon, chồn bạc ửng, chồn sương, vật nuôi trong nhà, thú nuôi, thỏ, chuột nhắt, chuột thường, chuột lang hoặc bộ gặm nhấm khác, hải cẩu, cá voi và các loài tương tự. Các sinh vật đích này là để dẫn chứng, và không được xem xét để giới hạn sáng chế và các ứng dụng của sáng chế.

Cũng được dự tính rằng thực vật chứa HA thể khảm theo một vài phương án thực hiện của sáng chế, hoặc biểu hiện VLP bao gồm HA thể khảm theo một vài phương án thực hiện theo sáng chế, có thể được sử dụng cho đối tượng hoặc sinh vật đích, theo các cách khác nhau phụ thuộc vào nhu cầu và hoàn cảnh. Ví dụ, HA thể khảm thu được từ thực vật có thể được chiết trước khi sử dụng nó dưới dạng thô, hoặc được tinh sạch một phần hoặc dạng được tinh sạch. Nếu HA thể khảm được tinh sạch ít nhất một phần, sau đó nó có thể được sản xuất trong các cây ăn được hoặc không ăn được. Hơn nữa, nếu HA thể khảm được sử dụng theo đường miệng, mô thực vật có thể được thu và cung cấp trực tiếp cho đối tượng, hoặc mô thu được có thể được làm khô trước khi cho ăn, hoặc động vật được cho ăn trên cây mà không cần thực hiện thu hoạch trước. Cũng được xem xét trong phạm vi của sáng chế với các mô cây đã thu hoạch được cung cấp làm nguồn thực phẩm dự trữ trong chăn nuôi gia súc. Nếu mô cây được cung cấp cho động vật một ít hoặc không xử lý thêm ưu tiên mô cây được sử dụng có thể ăn được.

Sự bất hoạt gen sau phiên mã (PTGS) có thể liên quan đến sự biểu hiện hạn chế của các gen chuyển trong thực vật, và biểu hiện đồng thời của tác nhân ức chế sự bất hoạt virus Y trên khoai tây (HcPro) có thể được sử dụng để chống lại sự thoái hóa đặc trưng của các mARN gen chuyển (Brigneti et al., 1998). Các tác nhân ức chế lần lượt sự bất

hoạt đã được biết trong lĩnh vực kỹ thuật và có thể được sử dụng như được mô tả ở đây (Chiba et al., 2006, *Virology* 346:7-14; được kết hợp ở đây để tham khảo), ví dụ nhưng không giới hạn với, TEV-p1/HC-Pro (virut-p1/HC-Pro khắc axit trên cây thuốc lá), BYV - p21, p19 của virut chùn ngọn tàn lụi cà chua (TBSV p19), protein vỏ capsit của virut làm quần lá cà chua (TCV -CP), 2b của virut khảm dưa chuột; CMV-2b), p25 của virut X trên khoai tây (PVX-p25), p11 của virut M trên khoai tây (PVM-p11), p11 của virut S trên khoai tây (PVS-p11), p16 của virut gây khô héo cây việt quất, (BScV -p16), p23 của Virut tàn lụi cam quýt (CTV-p23), p24 của virut-2 liên quan đến sâu cuốn lá nhỏ, (GLRaV-2 p24), p10 của virut A trên cây nho, (GVA-p10), p14 của virut B trên cây nho (GBV-p14), p10 của virut tiềm ẩn ở cây Heracleum (HLV-p10), hoặc p16 của virut tiềm ẩn phổ biến ở cây tỏi (GCLV-p16). Do đó, gen ức chế sự bất hoạt, ví dụ, nhưng không giới hạn với, HcPro, TEV -p1/HC-Pro, BYV-p21, TBSV p19, TCV-CP, CMV-2b, PVX-p25, PVM-p11, PVS-p11, BScV-p16, CTV-p23, GLRaV-2 p24, GBV-p14, HLV-p10, GCLV-p16 hoặc GVA-p10, có thể được biểu hiện đồng thời cùng với trình tự axit nucleic mã hóa protein được quan tâm để bảo đảm mức sản xuất protein trong cây cao.

Hơn nữa, các VLP được tạo ra như được mô tả ở đây không bao gồm neuraminidaza (NA). Tuy nhiên, NA có thể được biểu hiện đồng thời với HA nên các VLP bao gồm HA và NA được mong muốn.

Do đó, sáng chế còn gồm một vectơ thích hợp bao gồm trình tự HA thể khảm thích hợp với việc sử dụng các hệ thống biểu hiện ổn định hoặc tạm thời. Thông tin di truyền có thể được tạo ra trong một hoặc nhiều hơn một sản phẩm thiết kế. Ví dụ, trình tự nucleotit mã hóa protein được quan tâm có thể được đưa vào trong một sản phẩm thiết kế, và trình tự nucleotit thứ hai mã hóa protein biến đổi sự glycosyl hóa của protein được quan tâm có thể được đưa vào bằng cách sử dụng sản phẩm thiết kế riêng. Các trình tự nucleotit này có thể sau đó được biểu hiện đồng thời trong thực vật. Tuy nhiên, sản phẩm thiết kế bao gồm trình tự nucleotit mã hóa cả protein được quan tâm và protein biến đổi sự glycosyl hóa đặc trưng với protein được quan tâm cũng có thể được sử dụng. Trong trường hợp này trình tự nucleotit sẽ bao gồm trình tự thứ nhất bao gồm trình tự axit nucleic thứ nhất mã hóa protein được quan tâm được liên kết có hiệu quả với trình tự khởi đầu phiên mã hoặc vùng điều hòa, và trình tự thứ hai bao gồm trình tự axit nucleic thứ hai mã hóa protein biến đổi sự glycosyl hóa đặc trưng với protein được quan tâm, trình tự thứ hai được liên kết có hiệu quả với trình tự khởi đầu phiên mã hoặc vùng điều hòa.

Thuật ngữ “được biểu hiện đồng thời” nghĩa là hai, hoặc nhiều hơn hai, trình tự nucleotit được biểu hiện đồng thời trong thực vật, và trong cùng một mô thực vật. Tuy nhiên, các trình tự nucleotit không cần phải biểu hiện chính xác đồng thời. Hơn nữa, hai hoặc nhiều trình tự nucleotit được biểu hiện theo cách để các sản phẩm được mã hóa có cơ hội để tác động lẫn nhau. Ví dụ, protein biến đổi sự glycosyl hóa của protein được quan tâm có thể được biểu hiện trước hoặc trong khoảng thời gian khi protein đang nói đến được biểu hiện để thể hiện sự cải biến glycosyl hóa của protein được quan tâm. Hai hoặc nhiều hơn hai trình tự nucleotit có thể được biểu hiện đồng thời sử dụng hệ thống biểu hiện tạm thời, tại đó hai hoặc nhiều trình tự được đưa vào trong thực vật đồng thời dưới các điều kiện mà cả hai trình tự được biểu hiện. Mặt khác, thực vật nền bao gồm một trong các trình tự nucleotit, ví dụ trình tự mã hóa protein làm biến đổi đặc tính glycosyl hóa của protein quan tâm, có thể được biến nạp, theo cách tạm thời hoặc ổn định, cùng với trình tự bổ sung mã hóa protein quan tâm. Trong trường hợp này, trình tự mã hóa protein biến đổi đặc trưng glycosyl hóa của protein được quan tâm có thể được biểu hiện trong mô mong muốn, trong giai đoạn phát triển mong muốn, hoặc sự biểu hiện nó có thể được gây ra bằng cách sử dụng trình tự khởi đầu phiên mã cảm ứng được, và trình tự mã hóa protein được quan tâm có thể được biểu hiện dưới các điều kiện tương tự và trong mô tương tự, để bảo đảm rằng các trình tự nucleotit được biểu hiện đồng thời.

Các sản phẩm thiết kế theo sáng chế có thể được đưa vào các tế bào thực vật sử dụng các plasmit Ti, plasmit Ri, các vectơ virus thực vật, sự biến nạp ADN trực tiếp, vi tiêm, kỹ thuật xung điện, thâm nhiễm, và tương tự. Để xem lại các kỹ thuật này tham khảo ví dụ Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academy Press, New York VIII, pp 421-463 (1988); Geierman and Corey, *Plant Molecular Biology*, xuất bản lần thứ 2. (1988); and Miki and Iyer, *Fundamentals of Gene Transfer in Plants*. Trong *Plant Metabolism*, xuất bản lần thứ hai. DT. Dennis, DH Turpin, DD Lefebvre, DB Layzell (eds), Addison-Wesley, Langmans Ltd. London, pp. 561-579 (1997). Các phương pháp khác gồm kỹ thuật hấp thụ ADN trực tiếp, sử dụng các liposom, kỹ thuật xung điện, ví dụ sử dụng các thể nguyên sinh, vi tiêm, vi đạn hoặc sợi tinh thể, và thâm nhiễm bằng chân không. Tham khảo, ví dụ, Bilang, et al. (Gene 100: 247-250, 1991), Neuhauser et al. (Mol. Gen. Genet. 228: 104-112, 1991), Guerche et al. (Plant Science 52: 111-116, 1987), Neuhauser et al. (Theor. Appl Genet. 75: 30-36, 1987), Klein et al., Nature 327: 70-73 (1987); Howell et al. (Science 208: 1265, 1980), Horsch et al. (Science 227: 1229-1231,

1985), DeBlock và cộng sự, Plant Physiology 91: 694-701, 1989), , Liu and Lomonossoff (J. Virol Meth, 105:343-348, 2002,), Đơn xin cấp Patent Mỹ Số 4945050; 5036006; 5100792; 6403865; 5625136, (tất cả được kết hợp ở đây để tham khảo).

Các phương pháp biểu hiện tạm thời có thể được sử dụng để biểu hiện các sản phẩm thiết kế theo sáng chế (tham khảo Liu và Lomonossoff, 2002, Journal of Virological Methods, 105:343-348; được kết hợp ở đây để tham khảo). Mặt khác, phương pháp biểu hiện tạm thời dựa vào chân không, như được biểu hiện bởi Kapila et al. 1997 Plant Science 122:101-108 (được kết hợp ở đây để tham khảo) có thể được sử dụng. Các phương pháp này có thể gồm, ví dụ, nhưng không giới hạn với, phương pháp chủng bệnh bằng vi khuẩn Agro hoặc thâm nhiễm Agro, tuy nhiên, các phương pháp tạm thời khác cũng có thể được sử dụng như được chỉ ra ở trên. Bằng phương pháp lây nhiễm Agro hoặc thâm nhiễm Agro, hỗn hợp *Agrobacteria* bao gồm axit nucleic mong muốn đi vào không gian nội bào của mô, ví dụ các lá, phần trên không của cây (gồm thân, lá và hoa), các phần khác của cây (thân, rễ, hoa), hoặc toàn bộ cây. Sau khi đi qua các biểu bì Agro-inoculation nhiễm và truyền các bản sao t-ADN vào trong các tế bào. t-ADN được phiên mã thuộc thể bổ sung và mRNA được dịch mã, dẫn đến việc sản xuất protein được quan tâm trong các tế bào được tiêm nhiễm, tuy nhiên, sự truyền t-ADN trong nhân chỉ là tạm thời.

Các VLP bao gồm HA thể khảm được tạo ra bởi sáng chế có thể được sử dụng kết hợp với các vacxin cúm hiện nay, để bổ sung vacxin, làm cho chúng hiệu quả hơn, và làm giảm liều dùng cần thiết. Như đã biết bởi người có chuyên môn trong lĩnh vực kỹ thuật này, vacxin có thể được đề cập đối với một hoặc nhiều hơn một virus cúm. Các ví dụ về các vacxin thích hợp gồm, nhưng không giới hạn với, các vacxin có bán trên thị trường từ Sanofi-Pasteur, ID Biomedical, Merial, Sinovac, Chiron, Roche, MedImmune, GlaxoSmithKline, Novartis, Sanofi-Aventis, Serono, Shire Pharmaceuticals và tương tự.

Nếu cần, các VLP theo sáng chế có thể được trộn lẫn với tá dược thích hợp như đã biết bởi một người có chuyên môn trong lĩnh vực kỹ thuật. Hơn nữa, VLP có thể được sử dụng trong chế phẩm vacxin bao gồm liều lượng hữu hiệu của VLP để xử lý sinh vật đích, như được định rõ ở trên. Ngoài ra, VLP được sản xuất theo sáng chế có thể được kết hợp với các VLP thu được bằng cách sử dụng các protein cúm khác nhau, ví dụ, neuraminidaza (NA).

Do đó, sáng chế đề xuất phương pháp gây ra sự miễn dịch với sự nhiễm virus cúm trong động vật hoặc sinh vật đích bao gồm việc sử dụng một liều hữu hiệu vắc xin bao gồm một hoặc nhiều hơn một VLP. Vắc xin có thể được sử dụng theo đường miệng, trong da, trong mũi, trong cơ, trong bụng, trong tĩnh mạch, hoặc dưới da.

Các chế phẩm theo các phương án thực hiện khác nhau theo sáng chế có thể bao gồm các VLP của một hoặc nhiều chủng hoặc typ phụ cúm. “Hai hoặc nhiều hơn” đề cập đến hai, ba, bốn, năm, sáu, bảy, tám, chín, 10 hoặc nhiều hơn các chủng hoặc typ phụ. Các chủng hoặc typ phụ được thể hiện có thể là typ phụ đơn (ví dụ tất cả H1N1, hoặc tất cả H5N1), hoặc có thể là tổ hợp của các typ phụ. Các typ phụ và chủng minh họa gồm H5/Indo, H1/Bri, H1/NC, H3/Bri, B/Flo. Lựa chọn kết hợp giữa các chủng hoặc typ phụ có thể phụ thuộc vào vùng địa lý của các đối tượng có khả năng phơi nhiễm với bệnh cúm, các loài động vật hay tiếp xúc gần gũi với con người được chủng ngừa (ví dụ, các loài chim nước, các con vật nông nghiệp như lợn, v.v.) và các chủng mà chúng mang, được phơi nhiễm hoặc có khả năng phơi nhiễm, các dự đoán về sự lạc dòng kháng nguyên trong các typ phụ hoặc chủng, hoặc tổ hợp của các yếu tố này. Các ví dụ về các tổ hợp được sử dụng trong các năm trước trong các cơ sở dữ liệu được duy trì bởi Tổ chức Y tế thế giới (WHO) (tham khảo URL: who.int/csr/disease/influenza/vaccine_recommendations1/en).

Hai hoặc nhiều VLP có thể được biểu hiện riêng biệt, và các VLP được tinh sạch hoặc bán tinh sạch sau đó được kết hợp. Mặt khác, các VLP có thể được biểu hiện đồng thời trong cùng vật chủ, ví dụ cây, bộ phận của cây, hoặc tế bào cây. Các VLP có thể được kết hợp hoặc được tạo ra theo tỷ lệ mong muốn, ví dụ các tỷ lệ đương lượng, hoặc có thể được kết hợp theo cách mà một typ phụ hoặc chủng chủ yếu bao gồm các VLP trong chế phẩm.

Do đó, sáng chế đề xuất các chế phẩm bao gồm các VLP của hai hoặc nhiều chủng hoặc typ phụ.

Sáng chế cũng đề xuất vật phẩm sản xuất, bao gồm vật liệu bao gói và chế phẩm bao gồm VLP bao gồm HA thể khảm. Chế phẩm gồm tá dược được dùng hoặc sinh lý học, và vật liệu bọc có thể gồm nhãn chỉ ra các thành phần hoạt tính của chế phẩm (ví dụ VLP).

Kít thử bao gồm chế phẩm bao gồm axit nucleic mã hóa HA thể khảm được tạo ra trong bản mô tả, cùng với các chỉ dẫn sử dụng axit nucleic để sản xuất HA thể khảm,

hoặc các VLP bao gồm HA thể khảm cũng được tạo ra. Kít thử có thể hữu ích để sản xuất các VLP bao gồm HA thể khảm, và các chỉ dẫn có thể gồm, ví dụ, thông tin biểu hiện axit nucleic trong cây hoặc tế bào cây, các chỉ dẫn thu hoạch và thu các VLP từ cây hoặc mô cây.

Theo phương án thực hiện khác, kít thử để điều chế được phẩm, bao gồm VLP bao gồm HA thể khảm, cùng với các chỉ dẫn về việc sử dụng nó được tạo ra. Các chỉ dẫn có thể bao gồm một loạt các bước điều chế được phẩm, được phẩm hữu ích để gây ra đáp ứng miễn dịch trị liệu hoặc phòng ngừa trong đối tượng được sử dụng. Kít thử có thể bao gồm thêm các chỉ dẫn dùng riêng cho nồng độ liều dùng, các khoảng liều dùng, các phương pháp sử dụng được ưu tiên hoặc phương pháp tương tự.

Sáng chế cũng minh họa trong các ví dụ sau đây. Tuy nhiên được hiểu rằng các ví dụ này cho các mục đích minh họa, và không nên được sử dụng để giới hạn phạm vi của sáng chế theo cách bất kỳ.

Trình tự được mô tả trong bản mô tả được tóm tắt dưới đây.

SEQ ID NO	Mô tả	SEQ ID NO:	Mô tả
1	mỗi XmaI-pPlas.c	55	trình tự axit nucleic của BglII -PDI SP-H1/NC-SacI/StuI tổng hợp (thiếu TMD/CT)
2	mỗi SacI-ATG-pPlas.r	56	trình tự axit nucleic KpnI - H1/NC TmD/CT-SacI/StuI bao gồm trình tự mã hóa C-ter H1 (A/New Caledonia/20/99 (H1N1) gồm TmD và Ctail từ vị trí KpnI đến codon kết thúc (bị chặn tại đầu 3' nhờ vị trí SacI/StuI kép).
3	mỗi SacI-PlasTer.c	57	trình tự axit nucleic của protein disulfua isomeraza Nucleotit 32-103 mã hóa peptit tín hiệu PDI
4	mỗi EcoRI-PlasTer.r	58	trình tự axit nucleic của trình tự khởi đầu phiên mã plasto - PDI SP-BglI
5	mỗi Plasto-443c	59	trình tự axit nucleic của sản phẩm thiết kế 540 bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cổ linh lăng và 5' UTR, trình tự mã hóa peptit tín hiệu từ PDI và dạng H1 A/New Caledonia/20/99 (H1N1), plastoxyanin ở cổ linh lăng 3' UTR và trình tự kết thúc. H1 từ trình tự mã hóa A/New Caledonia/20/1999

			được gạch chân.
6	mỗi SpHA(Ind)-Plasto.r	60	trình tự axit nucleic của trình tự khởi đầu phiên mã DraI - Plasto bao gồm vùng mã hóa H1 hoàn thiện (A/Brisbane/59/07 (H1N1)) (gồm peptit tín hiệu và codon kết thúc) bị chặn, tại đầu 5' bởi trình tự gen plastoxyanin ở cở linh lăng tương ứng với 84 nucleotit thứ nhất ngược chiều với ATG ban đầu.
7	mỗi Plasto-SpHA	61	trình tự axit nucleic của sản phẩm thiết kế 774 bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cở linh lăng và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu có dạng H1 A/Brisbane/59/07 (H1N1), plastoxyanin ở cở linh lăng 3' UTR và trình tự kết thúc.
8	mỗi HA(Ind)-Sac.r	62	trình tự axit nucleic của sản phẩm thiết kế 828, từ PacI (ngược chiều với trình tự khởi đầu phiên mã) đến AscI (cùng chiều với nhân tố kết thúc NOS). Máy giải trình tự CPMV HT 3' UTR được gạch dưới bởi ATG đột biến in đậm.
9	mỗi pBinPlus.2613c	63	trình tự axit nucleic của sản phẩm thiết kế 690
10	mỗi Mut-ATG115.r	64	trình tự axit nucleic của sản phẩm thiết kế 691
11	mỗi Mut-ATG161.c	65	trình tự axit nucleic của sản phẩm thiết kế 696
12	mỗi LC-C5-1.110r	66	trình tự axit nucleic của sản phẩm thiết kế 732 bao gồm trình tự khởi đầu phiên mã CaMV 35S, CPMV-HT 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu có dạng H1 A/Brisbane/59/07 (H1N1), aCPMV-HT 3' UTR và trình tự kết thúc. Trình tự mã hóa H1/Bri được gạch dưới.
13	mỗi E1 H1B-E1 H5I.r	67	trình tự axit nucleic của sản phẩm thiết kế trung gian 787
14	mỗi E1 H5N-E1 H1B.c	68	trình tự axit nucleic của sản phẩm thiết kế 733
15	mỗi E2 H5I-RB H1B.r	69	trình tự axit nucleic của sản phẩm thiết kế 734
16	mỗi RB H1B-E2 H5I.c	70	trình tự axit nucleic của trình tự khởi đầu phiên mã DraIII-plasto - H3/Bri-

			SacI bao gồm vùng mã hóa H3 hoàn thiện (A/Brisbane/10/07 (H3N2)) (gồm peptit tín hiệu và codon kết thúc) bị chặn, tại đầu 5' bởi trình tự gen plastoxyanin ở cở linh lăng tương ứng với 84 nucleotit thứ nhất ngược chiều với ATG ban đầu.
17	mỗi E1 H1B-F'1 H5I.r	71	trình tự axit nucleic của sản phẩm thiết kế 736
18	mỗi F'1 H5N-E1 H1B.c	72	trình tự axit nucleic của sản phẩm thiết kế 737
19	mỗi F'2 H5I-E2 H1B.r	73	trình tự axit nucleic của trình tự khởi đầu phiên mã DraIII-plasto - B/Flo-SacI
20	mỗi E2 H1B-F'2 H5I.c	74	trình tự axit nucleic của sản phẩm thiết kế 739
21	mỗi E1 H5I-E1 H1NC.r	75	trình tự axit nucleic của sản phẩm thiết kế 745
22	mỗi E1 H1NC-E1 H5I.c	76	trình tự axit nucleic của Msj1
23	mỗi E2 H1NC-RB H5I.r	77	trình tự axit nucleic của sản phẩm thiết kế 850
24	mỗi RB H5I-E2 H1NC.c	78	trình tự axit nucleic của sản phẩm thiết kế 860
25	mỗi HA-SacI.r	79	trình tự axit nucleic của sản phẩm thiết kế 870
26	mỗi ApaI-H1B.c	80	Trình tự axit amin bao gồm vùng phụ RB của H1/Bri, peptit tín hiệu H5/Indo, và phức hệ vùng gốc (SDC) bao gồm các vùng phụ H5/Indo F'1, E1, E2, F'2 và F.
27	mỗi StuI-H1B.r	81	trình tự axit amin bao gồm phức hệ vùng đầu H1/Bri (HDC) bao gồm peptit tín hiệu E1, RB, E2, H5/Indo, và phức hệ vùng gốc H5/Indo (SDC) bao gồm vùng phụ H5/Indo F'1, F'2 và F.
28	mỗi SpPDI-H1B.c	82	trình tự axit amin bao gồm vùng phụ RB của H5/Indo, peptit tín hiệu PDI, và phức hệ vùng gốc H1/NC bao gồm F'1, E1, E2 và F'2.
29	mỗi SacI-H1B.r	83	trình tự axit amin bao gồm peptit tín hiệu PDI, vùng ngoài của H3 A/Brisbane/10/2007 và TMD/CT của H5 A/Indonesia/5/2005.
30	mỗi ApaI-SpPDI.c	84	trình tự axit nucleic bao gồm vùng ngoài của B/Florida/4/2006 và TMD/CT của H5 A/Indonesia/5/2005 được mã hóa

			bằng khung đọc mở trong sản phẩm thiết kế số 745.
31	mỗi ApaI-H5 (A-Indo).1c	85	trình tự axit nucleic của SacI - plastoxyanin 3'UTR
32	mỗi H5 (A-Indo)-StuI.1707r	86	trình tự nucleotit của peptit tín hiệu PDI SP-H1 A/California/04/09 protein disulfua isomeraza ở Cỏ linh lăng được gạch dưới và trình tự mã hóa H1 hoàn chỉnh được làm nổi bật nhờ in đậm.
33	mỗi H3B-SpPDI.r	87	trình tự axit amin của peptit tín hiệu protein disulfua isomeraza ở cỏ linh lăng PDI SP-H1 A/California/04/09 được gạch dưới
34	mỗi SpPDI-H3B.c	88	trình tự khởi đầu phiên mã 2X35S.
35	mỗi StuI-H3B.r	89	mỗi PacI -MCS-2X35S.c
36	mỗi TmD H5I-H3B.r	90	mỗi CPMV 5'UTR- 2X35S .r
37	mỗi H3B-TmD H5I.c	91	mỗi 2X35S -CPMV 5'UTR.c
38	mỗi HBF-SpPDI.r	92	mỗi Apal -M prot.r
39	mỗi SpPDI-HBF.c	93	trình tự axit nucleic của sản phẩm thiết kế 747. Trình tự mã hóa HA thể khảm được gạch dưới. Trình tự trình tự khởi đầu phiên mã 2X35S được chỉ ra nhờ in nghiêng.
40	mỗi StuI-HBF.r	94	trình tự axit amin của B-Florida
41	mỗi TmD H5I-B Flo.r	95	trình tự axit amin của B-Malaysia
42	mỗi B Flo-TmD H5I.c	96	trình tự axit amin của H1-Brisban
43	mỗi Hsp40Luz.1c	97	trình tự axit amin của H1-Sol.Isl.
44	mỗi Hsp40Luz-SacI.1272r	98	trình tự axit amin của H1-New Cal.
45	mỗi Hsp40Luz-Plasto.r	99	trình tự axit amin của H2-Singapore
46	mỗi Hsp70Ara.1c	100	trình tự axit amin của H3-Brisbane
47	mỗi Hsp70Ara-SacI.1956r	101	trình tự axit amin của H3A-WCN
48	mỗi Hsp70Ara-Plasto.r	102	trình tự axit amin của H5A-Anhui
49	mỗi supP19-plasto.r	103	trình tự axit amin của H5A-Vietnam
50	mỗi supP19-1c	104	trình tự axit amin của H5-Indo
51	mỗi SupP19-SacI.r		
52	trình tự axit nucleic vùng mã hóa H5 hoàn chỉnh (A/Indonesia/5/05 (H5N1)) (gồm peptit tín hiệu và codon kết thúc) được chặn, tại đầu 5'		

	bởi vị trí HindIII và, tại đầu 3', bởi vị trí SacI.		
53	trình tự axit nucleic của sản phẩm thiết kế 660 - catxet biểu hiện HA bao gồm trình tự khởi đầu phiên mã plastoxyanin ở cò linh lăng và 5' UTR, trình tự mã hóa chất ngưng kết hồng cầu có dạng H5 A/Indonesia/5/05 (H5N1), plastoxyanin ở cò linh lăng 3' UTR và trình tự kết thúc.		
54	trình tự axit nucleic của H1/NC gốc (thiếu TMD/CT)		

Phương pháp và vật liệu

Sự gắn các catxet biểu hiện HA

A- pCAMBIAPlasto

Tất cả các thao tác được tiến hành bằng cách sử dụng các phương thức sinh học phân tử nói chung của Sambrook and Russell (2001; được kết hợp ở đây để tham khảo). Bảng 1 thể hiện các môi oligonucleotit để gắn các catxet biểu hiện. Bước tách dòng thứ nhất gồm sự gắn plasmit thụ thể chứa các nhân tố điều hòa ngược dòng và xuôi dòng với gen plastoxyanin ở cò linh lăng. Trình tự khởi đầu phiên mã plastoxyanin và các trình tự 5'UTR được khuếch đại từ ADN của gen cò linh lăng bằng cách sử dụng các môi XmaI-pPlas.c (SEQ ID NO:1) và SacI-ATG-pPlas.r (SEQ ID NO:2) Sản phẩm khuếch đại thu được được phân cắt bằng XmaI và SacI và được gắn vào pCAMBIA2300 (Cambia, Canberra, Úc), đã được phân cắt từ trước bởi cùng các enzym, để tạo thành pCAMBIApromoPlasto. Tương tự, các trình tự 3'UTR và nhân tố kết thúc của gen plastoxyanin được phóng đại từ ADN hệ gen của cò linh lăng sử dụng các môi sau đây: SacI-PlasTer.c (SEQ ID NO:3) và EcoRI-PlasTer.r (SEQ ID NO:4) và sản phẩm được phân loại bởi SacI và EcoRI trước khi được gắn vào trong các cùng vị trí của pCAMBIApromoPlasto tạo ra pCAMBIAPlasto

B- Plasto- SP-H5 A/Indonesia/5/05 gốc (sản phẩm thiết kế số 660)

Đoạn mã hóa chất ngưng kết hồng cầu từ chủng cúm A/Indonesia/5/05 (H5N1; Số truy cập LANL ISDN125873) được tổng hợp bởi Epoch Biolabs (Sugar Land, TX, USA). Đoạn được tạo ra, chứa vùng mã hóa H5 hoàn chỉnh gồm peptit tín hiệu gốc bị chặn bởi vị trí HindIII ở đoạn ngược chiều trực tiếp của ATG khởi đầu, và vị trí SacI ngay phía xuôi dòng của codon (TAA) kết thúc, được thể hiện trong (SEQ ID NO: 25; Fig.17). Vùng mã hóa H5 được tách dòng vào catxet biểu hiện dựa trên plastoxyanin bằng phương pháp gắn dựa trên PCR được thể hiện trong Darveau et al. (1995). Tóm lại, sự phóng đại PCR lần đầu thu được bằng cách sử dụng các môi Plasto-443c (SEQ ID NO:5;) và SpHA(Ind)-Plasto.r (SEQ ID NO:6) và pCAMBIApromoPlasto làm khuôn. Tương tự, sự phóng đại thứ hai được tiến hành với các môi Plasto-SpHA (SEQ ID NO:7) và HA(Ind)-Sac.r (SEQ ID NO:8) với đoạn mã hóa H5 (SEQ ID NO:52; Fig.17) làm khuôn. Sản phẩm khuếch đại thu được từ cả hai phản ứng được trộn với nhau và hỗn hợp này được sử dụng làm khuôn cho phản ứng thứ ba (phản ứng liên hợp) bằng cách sử dụng Plasto-443c (SEQ ID NO:5) và HA(Ind)-Sac.r (SEQ ID NO:8) làm môi. Đoạn kết quả được phân hủy bởi BamHI (trong trình tự khởi đầu phiên mã plastoxyanin) và SacI (tại đầu 3' của đoạn) và được tách dòng thành pCAMBIAPlasto được phân loại trước bởi các enzym tương tự. Plasmit thu được, được đặt tên là 660, được thể hiện trên Fig.18 (SEQ ID NO: 53).

C- Plasto-PDI SP-H1 A/New Caledonia/20/99 (sản phẩm thiết kế số 540)

Cấu trúc khung đọc mở từ gen H1 của chủng cúm A/New Caledonia/20/99 (H1N1) được tổng hợp trong hai đoạn (Plant Biotechnology Institute, National Research Council, Saskatoon, Canada). Đoạn thứ nhất được tổng hợp tương ứng với trình tự mã hóa H1 kiểu đại (Ngân hàng gen số gia nhập AY289929; SEQ ID NO: 54; Fig.19) mà không có trình tự mã hóa peptit tín hiệu ở đầu 5' và trình tự mã hóa màng vận chuyển ở đầu 3'. Đầu 5' của đoạn gồm nucleotit cuối cùng mã hóa PDISP (gồm vị trí giới hạn BglIII) và vị trí SacI/StuI kép được bổ sung ngay phía xuôi chiều của codon kết thúc tại đầu 3' của đoạn, để thu được SEQ ID NO: 55 (Fig.20). Đoạn thứ hai mã hóa đầu tận cùng C của protein H1 (bao gồm màng vận chuyển và đuôi tế bào chất) từ vị trí KpnI đến codon kết thúc, và bị chặn tại đầu 3' bởi các vị trí giới hạn SacI và StuI cũng được tổng hợp (SEQ ID NO. 56; Fig.21).

Đoạn H1 thứ nhất được phân cắt bởi BglIII và SacI và được tách dòng thành các vị trí tương tự của vectơ nhị phân (pCAMBIAPlasto) chứa trình tự khởi đầu phiên mã plastoxyanin và 5'UTR được dung hợp với peptit tín hiệu của gen protein disulfua

isomeraza (PDI) ở cở linh lăng (các nucleotit 32-103; Số gia nhập Z11499; SEQ ID NO: 57; Fig.22) thu được gen khảm PDI-H1 xuôi chiều với các nhân tố điều hòa plastoxyanin. Trình tự của catxet chứa plastoxyanin, chứa trình tự khởi đầu phiên mã và peptit tín hiệu PDI lên đến vị trí giới hạn BglII và nhân tố kết thúc plastoxyanin ngược chiều với vị trí SacI, được thể hiện trong SEQ ID NO. 58 (Fig.23). Sự thêm đầu tận cùng C của vùng mã hóa H1 (mã hóa vùng xuyên màng và đuôi tế bào chất) thu được bằng cách chèn đoạn tổng hợp (SEQ ID NO. 56; Fig.21) được phân cắt trước đó bằng KpnI và SacI, vào trong plasmit biểu hiện H1. Sản phẩm thiết kế thu được, được đặt tên là 540, được thể hiện trong SEQ ID NO. 59 (Fig.24).

D- Plasto- SP-H1 A/Brisbane/59/07 gốc (sản phẩm thiết kế số 774)

Catxet biểu hiện số 774, tiến hành biểu hiện H1 từ A/Brisbane/59/07, được tập hợp như sau. Đoạn tổng hợp được tổng hợp bao gồm trình tự mã hóa chất ngưng kết hồng cầu hoàn thiện (từ ATG đến kết thúc) bị chặn tại đầu 3' bởi trình tự gen plastoxyanin ở cở linh lăng tương ứng với 84 nucleotit đầu tiên ngược chiều với plastoxyanin ATG bắt đầu bằng vị trí giới hạn DraIII. Các đoạn tổng hợp cũng bao gồm vị trí SacI xuôi chiều với codon kết thúc.

Đoạn tổng hợp được tổng hợp bởi Top Gene Technologies (Montreal, QC, Canada). Đoạn được tổng hợp được thể hiện trong SEQ ID NO. 60 (Fig.25). Để tập hợp catxet biểu hiện hoàn thiện, đoạn tổng hợp được phân loại bởi DraIII và SacI và được tách dòng thành pCAMBIAPlasto được phân hủy trước bởi cùng các enzym tạo ra sản phẩm thiết kế 774 (SEQ ID NO. 61; Fig.26).

E- CPMV HT-LC C51 (sản phẩm thiết kế số 828)

Các catxet biểu hiện CPMV-*HT* sử dụng trình tự khởi đầu phiên mã 35S để kiểm soát sự biểu hiện mARN bao gồm trình tự mã hóa được quan tâm bị chặn, tại đầu 5', bởi các nucleotit 1-512 từ virus gây bệnh khảm cây đậu đũa (CPMV) ARN2 có ATG được biến nạp tại các vị trí 115 và 161 và tại đầu 3', bởi các nucleotit 3330-3481 từ CPMV ARN2 (tương ứng với 3' UTR) sau đó bởi tác nhân kết thúc NOS. Plasmit pBD-C5-1LC, (Sainsbury et al. 2008; Plant Biotechnology Journal 6: 82-92 và Đơn công bố PCT WO 2007/135480), được sử dụng để gắn các catxet biểu hiện Chất ngưng kết hồng cầu dựa trên CPMV-*HT*. Sự biến nạp các ATG tại vị trí 115 và 161 của CPMV ARN2 được tiến hành bằng cách sử dụng phương pháp gắn dựa trên PCR được thể hiện trong Darveau et

al. (Methods in Neuroscience 26: 77-85 (1995)). Hai PCR riêng biệt được tiến hành bằng cách sử dụng pBD-C5-1LC làm khuôn. Các môi cho quá trình khuếch đại đầu tiên là pBinPlus.2613c (SEQ ID NO: 9) và Mut-ATG115.r (SEQ ID NO: 10). Các môi cho quá trình khuếch đại thứ hai là Mut-ATG161.c (SEQ ID NO: 11) và LC-C5-1.110r (SEQ ID NO: 12). Hai đoạn thu được sau đó được trộn và sử dụng làm khuôn cho quá trình khuếch đại thứ ba bằng cách sử dụng pBinPlus.2613c (SEQ ID NO: 9) và LC-C5-1.110r (SEQ ID NO: 12) làm các môi. Đoạn thu được được phân cắt bằng PacI và ApaI và được tách dòng thành pBD-C5-1LC được phân cắt bởi cùng các enzym. Sản phẩm thiết kế được tạo ra, được đặt tên là 828, được thể hiện trên Fig.27 (SEQ ID NO: 62).

Vùng liên kết thụ thể F- H1 A/Brisbane/59/07 (RB) trong đoạn chính H5 A/Indonesia/5/05 (sản phẩm thiết kế số 690)

HA thể khảm được tạo thành bằng cách thay thế vùng RB trong H5 A/Indonesia/5/05 với vùng H1 A/Brisbane/59/07 sử dụng phương pháp gắn dựa trên PCR được thể hiện trong Darveau et al. (Methods in Neuroscience 26: 77-85(1995)). Trong chu kỳ PCR thứ hai, đoạn trình tự khởi đầu phiên mã plastoxyanin được dung hợp bởi peptit tín hiệu gốc, các vùng F'1 và E1 của H5 A/Indonesia/5/05 được phóng đại bằng cách sử dụng các môi Plasto-443c (SEQ ID NO: 5) và E1 H1B-E1 H5I.r (SEQ ID NO:13) với sản phẩm thiết kế số 660 (SEQ ID NO:53, Fig.18) dưới dạng khuôn. Đoạn thứ hai, bao gồm trình tự mã hóa vùng H1 A/Brisbane/59/07 RB, được phóng đại bởi các môi E1 H5N-E1 H1B.c (SEQ ID NO:14) và E2 H5I-RB H1B.r (SEQ ID NO:15) sử dụng sản phẩm thiết kế số 774 (SEQ ID NO:61; Fig.26) làm khuôn. Đoạn thứ ba bao gồm E2, F'2, F, vùng xuyên màng và tế bào chất từ H5 A/Indonesia/5/05 được phóng đại bằng cách sử dụng các môi RB H1B-E2 H5I.c (SEQ ID NO:16) và HA(Ind)-SacI.r (SEQ ID NO:8) với sản phẩm thiết kế số 660 (SEQ ID NO: 53; Fig.18) làm khuôn. Các sản phẩm khuếch đại sau đó được trộn và sử dụng làm khuôn cho chu kỳ khuếch đại thứ hai (phản ứng ghép nối) với các môi Plasto-443c (SEQ ID NO:5) và HA(Ind)-SacI.r (SEQ ID NO:8). Đoạn thu được được phân hủy bởi BamHI (trong trình tự khởi đầu phiên mã plastoxyanin) và SacI (sau codon kết thúc) và được tách dòng thành sản phẩm thiết kế số 660 (SEQ ID NO: 53; Fig.18), được phân hủy trước bởi các enzym cắt tương tự tạo ra sản phẩm thiết kế số 690 (SEQ ID NO: 63). Sản phẩm thiết kế được thể hiện trên Fig.28.

G- H1 A/Brisbane/59/07 esteraza và các vùng liên kết thụ thể (E1-RB-E2) trong đoạn chính H5 A/Indonesia/5/05 (sản phẩm thiết kế số 691)

HA thể khảm được gắn bằng cách thay thế các vùng E1-RB-E2 trong H5 A/Indonesia/5/05 với các vùng chứa H1 A/Brisbane/59/07 sử dụng phương pháp gắn dựa trên PCR được thể hiện trong Darveau et al. (Methods in Neuroscience 26: 77-85(1995)). Trong chu kỳ PCR thứ nhất, đoạn trình tự khởi đầu phiên mã plastoxyanin được dung hợp với peptit tín hiệu gốc và vùng F'1 của H5 A/Indonesia/5/05 được phóng đại sử dụng môi Plasto-443c (SEQ ID NO:5) và E1 H1B-F'1 H5I.r (SEQ ID NO: 17) với sản phẩm thiết kế số 660 (SEQ ID NO: 53; Fig.18) làm khuôn. Tương tự, hai đoạn khác nhau được phóng đại. Đoạn thứ hai, chứa trình tự mã hóa vùng H1 A/Brisbane/59/07 E1-RB-E2, được phóng đại bởi các môi F'1 H5N-E1 H1B.c (SEQ ID NO: 18) và F'2 H5I-E2 H1B.r (SEQ ID NO: 19) sử dụng sản phẩm thiết kế số 774 (SEQ ID NO:61; Fig.26) làm khuôn. Với đoạn thứ ba, F'2, F, vùng xuyên màng và tế bào chất từ H5 A/Indonesia/5/05 được phóng đại bằng cách sử dụng các môi E2 H1B-F'2 H5I.c (SEQ ID NO: 20) và HA(Ind)-Sac.r (SEQ ID NO: 8) với sản phẩm thiết kế số 660 (SEQ ID NO: 53; Fig.18) làm khuôn. Các sản phẩm khuếch đại sau đó được trộn và sử dụng làm khuôn cho chu kỳ khuếch đại thứ hai (phản ứng ghép nối) với các môi Plasto-443c (SEQ ID NO:5) và HA(Ind)-SacI.r (SEQ ID NO: 8). Đoạn kết quả được phân hủy bởi BamHI (trong trình tự khởi đầu phiên mã plastoxyanin) và SacI (sau codon kết thúc) và được tách dòng thành sản phẩm thiết kế số 660 (SEQ ID NO: 53; Fig.18), được phân hủy trước bởi các enzym giới hạn giống nhau tạo ra sản phẩm thiết kế số 691 (SEQ ID NO: 64). Sản phẩm thiết kế được thể hiện trên Fig.29.

Vùng liên kết thụ thể H- H5 A/Indonesia/5/05 (RB) trong đoạn chính H1 A/New Caledonia/20/99 (sản phẩm thiết kế số 696)

HA thể khảm được tạo thành bằng cách thay thế vùng RB trong H1 A/New Caledonia/20/99 với vùng H5 A/Indonesia/5/05 sử dụng phương pháp gắn được thể hiện trong Darveau et al. (Methods in Neuroscience 26: 77-85(1995)). Trong chu kỳ PCR đầu tiên, đoạn trình tự khởi đầu phiên mã plastoxyanin được dung hợp với peptit tín hiệu của protein disulfua isomeraza ở cở linh lăng (PDISP; Số gia nhập Z11499; các nucleotit 32-103 có SEQ ID NO: 57; Fig.22), các vùng F'1 và E1 của H1 A/New Caledonia/20/99 được phóng đại bằng cách sử dụng các môi Plasto-443c (SEQ ID NO: 5) và E1 H5I-E1 H1NC.r (SEQ ID NO: 21) với sản phẩm thiết kế số 540 (SEQ ID NO: 59; Fig.24) làm khuôn. Đoạn thứ hai, bao gồm trình tự mã hóa vùng H5 A/Indonesia/5/05 RB, được phóng đại bởi môi E1 H1NC-E1 H5I.c (SEQ ID NO: 22) và E2 H1NC-RB H5I.r (SEQ ID

NO: 23) sử dụng sản phẩm thiết kế số 660 (SEQ ID NO: 53; Fig.18) làm khuôn. Đoạn thứ ba bao gồm vùng vận chuyển E2, F'2, F, và vùng tế bào chất từ H1 A/New Caledonia/20/99 được phóng đại bằng cách sử dụng các môi RB H5I-E2 H1NC.c (SEQ ID NO: 24) và HA-SacI.r (SEQ ID NO: 25) với sản phẩm thiết kế số 540 (SEQ ID NO: 59; Fig.24) làm khuôn. Các sản phẩm khuếch đại sau đó được trộn và sử dụng làm khuôn cho chu kỳ khuếch đại thứ hai (phản ứng ghép nối) với các môi Plasto-443c (SEQ ID NO: 5) và HA-SacI.r (SEQ ID NO: 25). Đoạn thu được được phân cắt bởi BglIII và SacI và được tách dòng thành trong sản phẩm thiết kế số 540 (SEQ ID NO: 59; Fig.24) được phân hủy trước bởi các enzym giới hạn giống nhau to give sản phẩm thiết kế số 696 (SEQ ID NO: 65). Sản phẩm thiết kế này được thể hiện trên Fig.30.

Sự liên hợp I- của H1 A/Brisbane/59/2007 trong catxet biểu hiện CPMV-HT (sản phẩm thiết kế số 732).

Trình tự mã hóa HA từ H1 A/Brisbane/59/2007 được tách dòng thành CPMV-HT như sau. Các vị trí giới hạn ApaI (ngay phía xuôi chiều của ATG) và StuI (ngay phía xuôi chiều của codon kết thúc) được thêm vào trình tự mã hóa chất ngưng kết hồng cầu bằng cách tiến hành quá trình khuếch đại PCR bằng các môi ApaI-H1B.c (SEQ ID NO: 26) và StuI-H1B.r (SEQ ID NO: 27) sử dụng sản phẩm thiết kế số 774 (SEQ ID NO: 61; Fig.26) làm khuôn. Đoạn thu được được phân cắt bởi các enzym giới hạn ApaI và StuI và được tách dòng thành sản phẩm thiết kế số 828 (SEQ ID NO: 62; Fig.27) được phân cắt bởi cùng các enzym. Catxet thu được được đặt tên là sản phẩm thiết kế số 732 (SEQ ID NO: 66; Fig.31).

Sự liên hợp J của SpPDI-H1 A/Brisbane/59/2007 ở catxet biểu hiện CPMV-HT (sản phẩm thiết kế số 733).

Trình tự mã hóa peptit tín hiệu của protein disulfua isomeraza ở cỏ linh lăng (PDISP; các nucleotit 32-103 có SEQ ID NO: 57 Fig.22; Số gia nhập Z11499) được dung hợp với trình tự mã hóa HA0 của H1 từ A/Brisbane/59/2007, và đoạn thu được được tách dòng thành CPMV-HT như sau. Trình tự mã hóa H1 được khuếch đại bằng các môi SpPDI-H1B.c (SEQ ID NO: 28) và SacI-H1B.r (SEQ ID NO: 29) sử dụng sản phẩm thiết kế 774 (SEQ ID NO: 61; Fig.26) làm khuôn. Đoạn thu được gồm trình tự mã hóa H1 bị chặn, tại đầu 5', bởi các nucleotit sau cùng mã hóa PDISP (gồm vị trí giới hạn BglIII) và, tại đầu 3', bởi vị trí giới hạn SacI. Đoạn thu được được phân cắt bởi các enzym giới hạn

ApaI và StuI và được tách dòng thành sản phẩm thiết kế số 540 (SEQ ID NO: 59; Fig.24) được phân hủy trước bởi các enzym giới hạn giống nhau. Trình tự mã hóa của catxet trung gian, được gọi tên là sản phẩm thiết kế số 787 (SEQ ID NO: 67), được thể hiện trên Fig.32. Các vị trí giới hạn ApaI (ngay phía xuôi chiều của ATG) và StuI (ngay phía xuôi dòng của codon kết thúc) được thêm vào trình tự mã hóa chất ngưng kết hồng cầu bằng cách tiến hành quá trình khuếch đại PCR với các môi ApaI-SpPDI.c (SEQ ID NO: 30) và StuI-H1B.r (SEQ ID NO: 27) sử dụng sản phẩm thiết kế số 787 (SEQ ID NO: 67; Fig.32) làm khuôn. Đoạn thu được được phân cắt bằng các enzym giới hạn ApaI và StuI và được tách dòng thành sản phẩm thiết kế số 828 (SEQ ID NO: 62; Fig.27) được phân cắt bởi cùng một enzym. Catxet thu được được đặt tên là sản phẩm thiết kế số 733 (SEQ ID NO: 68; Fig.33).

K- Sự gắn vùng liên kết thụ thể H1 A/Brisbane/59/07 (RB) trong đoạn chính H5 A/Indonesia/5/05 trong catxet biểu hiện CPMV-HT (sản phẩm thiết kế số 734).

Trình tự mã hóa HA thể khảm gồm vùng RB từ H1 A/Brisbane/59/07 trong đoạn chính H5 A/Indonesia/5/05 được tách dòng thành CPMV-HT như sau. Các vị trí giới hạn ApaI (ngay phía xuôi chiều của ATG) và StuI (ngay phía xuôi chiều của codon kết thúc) được thêm vào trình tự mã hóa chất ngưng kết hồng cầu thể khảm bằng cách tiến hành quá trình khuếch đại PCR bởi các môi ApaI-H5 (A-Indo).1c (SEQ ID NO: 31) và H5 (A-Indo)-StuI.1707r (SEQ ID NO: 32) sử dụng sản phẩm thiết kế số 690 (SEQ ID NO: 63; Fig.28) làm khuôn. Đoạn thu được được phân cắt bằng các enzym giới hạn ApaI và StuI và được tách dòng thành sản phẩm thiết kế số 828 (SEQ ID NO: 62; Fig.27) được phân cắt bởi cùng một enzym. Catxet thu được được đặt tên là sản phẩm thiết kế số 734 (SEQ ID NO: 69; Fig.34).

Sự liên hợp L của SpPDI-H3 A/Brisbane/10/2007 trong catxet biểu hiện CPMV-HT (sản phẩm thiết kế số 736).

Trình tự mã hóa peptit tín hiệu PDI cở linh lăng được dung hợp với HA0 từ H3 A/Brisbane/10/2007 được tách dòng thành CPMV-HT như sau. Đầu tiên, đoạn tổng hợp được tổng hợp bao gồm trình tự mã hóa chất ngưng kết hồng cầu hoàn thiện (từ ATG đến kết thúc) được chặn bởi trình tự gen plastoxyanin ở cở linh lăng tương ứng với 84 nucleotit (bắt đầu bằng vị trí giới hạn DraIII) ngược chiều với plastoxyanin ATG. Đoạn tổng hợp cũng bao gồm vị trí SacI ngay sau codon kết thúc. Đoạn tổng hợp được tổng hợp

bởi Top Gene Technologies (Montreal, QC, Canada). Đoạn được tổng hợp được thể hiện trong SEQ ID NO: 70 (Fig.35) và được sử dụng làm khuôn cho sự gắn dựa trên PCR.

Thứ hai, peptit tín hiệu của protein disulfua isomeraza (PDISP) ở cỏ linh lăng (các nucleotit 32-103; Số gia nhập Z11499; SEQ ID NO: 57; Fig.22) được liên kết với trình tự mã hóa HA0 của H3 từ A/Brisbane/10/2007 cùng với vị trí giới hạn ApaI ngay cùng chiều với vị trí giới hạn ATG và StuI ngược chiều với codon kết thúc như sau. PDISP được liên kết với trình tự mã hóa H3 bởi phương pháp gắn dựa trên PCR được thể hiện trong Darveau et al. (Methods in Neuroscience 26: 77-85(1995)). Trong chu trình PCR đầu tiên, một đoạn của trình tự khởi đầu phiên mã plastoxyanin được khuếch đại bằng cách sử dụng các môi Plasto-443c (SEQ ID NO: 30) và H3B-SpPDI.r (SEQ ID NO: 33) với sản phẩm thiết kế số 540 (SEQ ID NO: 59; Fig.24) làm khuôn. Tương tự, đoạn khác chứa một phần của trình tự mã hóa của H3 A/Brisbane/10/2007 (từ codon 17 đến codon kết thúc) được khuếch đại bởi các môi SpPDI-H3B.c (SEQ ID NO: 34) và StuI-H3B.r (SEQ ID NO: 35) sử dụng đoạn được tổng hợp trước (SEQ ID NO. 70; Fig.35) làm khuôn. Các sản phẩm phóng đại sau đó được trộn và được sử dụng làm khuôn cho chu kỳ phóng đại thứ hai (phản ứng ghép nối) với các môi ApaI-SpPDI.c (SEQ ID NO: 30) và StuI-H3B.r (SEQ ID NO: 35). Đoạn thu được được phân cắt bởi các enzym giới hạn ApaI và StuI và được tách dòng thành sản phẩm thiết kế số 828 (SEQ ID NO: 62; Fig.27) được phân cắt bởi cùng một enzym. Catxet thu được được đặt tên là sản phẩm thiết kế số 736 (SEQ ID NO: 71; Fig.36).

M- Sự gắn SpPDI-H3 A/Brisbane/10/2007 thể khảm (vùng ngoài) + H5 A/Indonesia/5/2005 (TmD + đuôi Xyto) trong catxet biểu hiện CPMV-HT (sản phẩm thiết kế số 737).

Trình tự mã hóa peptit tín hiệu PDI cỏ linh lăng được dung hợp với vùng ngoài của H3 A/Brisbane/10/2007 và với các vùng xuyên màng và tế bào chất của H5 A/Indonesia/5/2005 được tách dòng thành CPMV-HT như sau. Trình tự mã hóa PDISP-H3 được dung hợp với vùng xuyên màng H5 bởi phương pháp gắn dựa trên PCR được thể hiện trong Darveau et al. (Methods in Neuroscience 26: 77-85(1995)). Trong chu kỳ đầu của PCR, đoạn bao gồm peptit tín hiệu PDISP và vùng ngoài từ H3 Brisbane được tạo ra bằng cách phóng đại (với vị trí giới hạn ApaI ngược chiều với ATG ban đầu PDISP ATG) sử dụng các môi ApaI-SpPDI.c (SEQ ID NO: 30) và TmD H5I-H3B.r (SEQ ID NO: 36) với sản phẩm thiết kế số 736 (SEQ ID NO: 71; Fig.36) làm khuôn. Tương tự, đoạn khác

chứa các vùng xuyên màng và tế bào chất của H5 Indonesia được phóng đại bởi các môi H3B-TmD H5I.c (SEQ ID NO: 37) và H5 (A-Indo)-StuI.1707r (SEQ ID NO: 32) sử dụng sản phẩm thiết kế số 660 (SEQ ID NO. 53; Fig.18) làm khuôn. Các sản phẩm phóng đại được trộn và sử dụng làm khuôn cho chu kỳ phóng đại thứ hai (phản ứng ghép nối) bởi các môi ApaI-SpPDI.c (SEQ ID NO: 30) và H5 (A-Indo)-StuI.1707r (SEQ ID NO: 32). Đoạn thu được được phân cắt bởi các enzym giới hạn ApaI và StuI và được tách dòng thành sản phẩm thiết kế số 828 (SEQ ID NO: 62; Fig.27) được phân cắt bởi cùng một enzym. Catxet thu được được đặt tên là sản phẩm thiết kế số 737 (SEQ ID NO: 72; Fig.37).

N- Sự gắn SpPDI-HA B/Florida/4/2006 trong catxet biểu hiện CPMV-HT (sản phẩm thiết kế số 739).

Trình tự mã hóa peptit tín hiệu PDI ở cỏ linh lăng được dung hợp với HA0 từ HA B/Florida/4/2006 được tách dòng thành CPMV-HT như sau. Đầu tiên, đoạn tổng hợp được tổng hợp bao gồm trình tự mã hóa chất ngưng kết hồng cầu hoàn thiện (từ ATG đến kết thúc) được chặn bởi trình tự gen plastoxyanin ở cỏ linh lăng tương ứng với 84 nucleotit (bắt đầu bằng vị trí giới hạn DraIII) ngược chiều với plastoxyanin ATG. Đoạn tổng hợp cũng bao gồm vị trí giới hạn SacI ngay sau codon kết thúc. Đoạn tổng hợp được tổng hợp bởi Epoch Biolabs (Sugar Land, Texas, USA). Đoạn đã tổng hợp được thể hiện trong SEQ ID NO: 73 (Fig.38) và được sử dụng làm khuôn cho sự gắn dựa trên PCR.

Thứ hai, peptit tín hiệu protein disulfua isomeraza (PDISP) ở cỏ linh lăng (các nucleotit 32-103 có SEQ ID NO: 57; Fig.22; Số gia nhập Z11499) được liên kết với trình tự mã hóa HA0 HA từ B/Florida/4/2006 cùng với vị trí giới hạn ApaI ngay ngược chiều với ATG và vị trí giới hạn StuI condon kết thúc ngược chiều như sau. PDISP được liên kết với trình tự mã hóa HA bởi phương pháp gắn dựa trên PCR được thể hiện trong Darveau et al. (Methods in Neuroscience 26: 77-85(1995)). Trong chu trình PCR đầu tiên, một đoạn của trình tự khởi đầu phiên mã plastoxyanin được khuếch đại bằng cách sử dụng các môi Plasto-443c (SEQ ID NO: 30) và HBF-SpPDI.r (SEQ ID NO: 38) với sản phẩm thiết kế số 540 (SEQ ID NO: 59; Fig.24) làm khuôn. Tương tự, đoạn khác chứa một phần của trình tự mã hóa HA từ B/Florida/4/2006 (từ codon 16 đến codon kết thúc) được phóng đại bởi các môi SpPDI-HBF.c (SEQ ID NO: 39) và StuI-HBF.r (SEQ ID NO: 40) sử dụng đoạn được tổng hợp trước (SEQ ID NO. 73; Fig.38) làm khuôn. Các sản phẩm phóng đại sau đó được trộn và sử dụng làm khuôn với chu kỳ phóng đại thứ hai (phản ứng ghép nối)

với các môi *ApaI*-SpPDI.c (SEQ ID NO: 30) và *StuI*-HBF.r (SEQ ID NO: 40). Đoạn thu được được phân cắt bởi các enzym giới hạn *ApaI* và *StuI* và được tách dòng thành sản phẩm thiết kế số 828 (SEQ ID NO: 62; Fig.27) được phân cắt bởi cùng một enzym. Catxet thu được được đặt tên là sản phẩm thiết kế số 739 (SEQ ID NO: 74; Fig.39).

O- Sự ghép SpPDI-HA B/Florida/4/2006 thể khảm (vùng ngoài) + H5 A/Indonesia/5/2005 (TmD + đuôi Xyto tail) trong catxet biểu hiện CPMV-HT (sản phẩm thiết kế số 745).

Trình tự mã hóa peptit tín hiệu PDI cổ linh lãg được dung hợp với vùng ngoài từ HA B/Florida/4/2006 và vùng xuyên màng và tế bào chất của H5 A/Indonesia/5/2005 được tách dòng thành CPMV-HT như sau. Vùng ngoài PDISP-B/Florida/4/2006 mã hóa trình tự được dung hợp với vùng xuyên màng và tế bào chất bằng phương pháp gắn dựa trên PCR được thể hiện trong Darveau et al. (Methods in Neuroscience 26: 77-85(1995)). Trong chu kỳ PCR thứ nhất, đoạn bao gồm peptit tín hiệu PDISP được dung hợp với vùng ngoài từ HA B/Florida/4/2006 được tạo ra bằng cách phóng đại sử dụng các môi *ApaI*-SpPDI.c (SEQ ID NO: 30) và TmD H5I-B Flo.r (SEQ ID NO: 41) với sản phẩm thiết kế số 739 (SEQ ID NO: 74; Fig.39) làm khuôn. Tương tự, đoạn khác chứa các vùng xuyên màng và tế bào chất H5 được phóng đại bởi các môi B Flo-TmD H5I.c (SEQ ID NO: 42) và H5 (A-Indo)-*StuI*.1707r (SEQ ID NO: 32) sử dụng sản phẩm thiết kế số 660 (SEQ ID NO: 53; Fig.18) làm khuôn. Sản phẩm phóng đại sau đó được trộn và sử dụng làm khuôn cho chu kỳ phóng đại thứ hai (phản ứng ghép nối) với các môi *ApaI*-SpPDI.c (SEQ ID NO: 30) và H5 (A-Indo)-*StuI*.1707r (SEQ ID NO: 32). Đoạn thu được được phân cắt bởi các enzym giới hạn *ApaI* và *StuI* và được tách dòng thành sản phẩm thiết kế số 828 (SEQ ID NO: 62; Fig.27) được phân cắt bởi cùng một enzym. Catxet thu được được đặt tên là sản phẩm thiết kế số 745 (SEQ ID NO: 75; Fig.40).

P-Sự gắn SpPDI-HA B/Florida/4/2006+ H5 A/Indonesia/5/2005 thể khảm (TmD + đầu Xyto) trong catxet biểu hiện 2X35S-CPMV-HT (sản phẩm thiết kế số 747).

Trình tự mã hóa peptit tín hiệu PDI cổ linh lãg được dung hợp với HA0 từ HA B/Florida/4/2006 và vùng xuyên màng và tế bào chất của H5 A/Indonesia/5/2005 được tách dòng thành 2X35S-CPMV-HT như sau. Sự chuyển đổi trình tự khởi đầu phiên mã được tiến hành bằng cách sử dụng phương pháp gắn dựa trên PCR được thể hiện trong Darveau et al. (Methods in Neuroscience 26: 77-85 (1995)). Đoạn thứ nhất chứa trình tự

khởi đầu phiên mã 2X35S (SEQ ID NO: 88; Fig.50A) được phóng đại bởi PCR bởi các môi **PacI**-MCS-2X35S.c (SEQ ID NO: 89) và CPMV 5'UTR-**2X35S**.r (SEQ ID NO: 90):

PacI-MCS-2X35S.c (SEQ ID NO:89)

AATTGTTAATTAAGTCGACAAGCTTGCATGCCTGCAGGTCAAC

CPMV 5'UTR-**2X35S**.r (SEQ ID NO:90)

TCAAAACCTATTAAGATTTTAATACCTCTCCAAATGAAATGAACTTCC

sử dụng plasmit chứa trình tự khởi đầu phiên mã 2X35S làm khuôn. Tương tự, PCR thứ hai được tiến hành bằng cách sử dụng các môi **2X35S**-CPMV 5'UTR.c (SEQ ID NO: 91) và **StuI**-HBF.r (SEQ ID NO: 92):

2X35S-CPMV 5'UTR.c (SEQ ID NO:91)

TTGGAGAGGTATTAATAATCTTAATAGGTTTTGATAAAAGCGAACGTGGG

Apal-M prot.r (SEQ ID NO:92)

TCTCCATGGGCCCCGACAAATTTGGGCAGAATATACAGAAGCTTA

sử dụng sản phẩm thiết kế 745 (SEQ ID NO 75; Fig.40) làm khuôn. Hai đoạn thu được sau đó được sử dụng làm khuôn với chu kỳ thứ hai của PCR (phản ứng ghép nối) với **PacI**-MCS-2X35S.c (SEQ ID NO:89 và **Apal**-M prot.r (SEQ ID NO:92) làm các khuôn. Đoạn thu được sau đó được phân cắt bằng **PacI** và **Apal** và được tách dòng thành trong sản phẩm thiết kế 745 (SEQ ID NO 75; Fig.74) được phân cắt bởi cùng các enzym giới hạn. Trình tự của vectơ trung gian, được đặt tên là 747 (SEQ ID NO:93), được thể hiện trên Fig.94.

Sự liên hợp của các catxet biểu hiện chaperon

Hai catxet biểu hiện protein sốc nhiệt (Hsp) được ghép nối. Trong catxet thứ nhất, biểu hiện của *Arabidopsis thaliana* (kiểu sinh thái Columbia) HSP70 xytosol (Athsp70-1 trong Lin et al. (2001) Cell Stress and Chaperones 6: 201-208) được kiểm soát bằng một trình tự khởi đầu phiên mã khả năng kết hợp các yếu tố Nitrit reductaza (Nir) ở cỏ linh lăng và các trình tự khởi đầu phiên mã plastoxyanin ở cỏ linh lăng (Nir/Plasto). Catxet thứ hai bao gồm vùng mã hóa của xytosol HSP40 ở cỏ linh lăng (MsJ1; Frugis và cộng sự. (1999) Plant Molecular Biology 40: 397-408) dưới sự kiểm soát của trình tự khởi đầu phiên mã khả năng Nir/Plasto cũng được liên hợp.

Một plasmit tiếp nhận chứa trình tự khởi đầu phiên mã Nitrit reductaza (Nir) ở cỏ linh lăng, gen chỉ thị GUS và tác nhân kết thúc NOS trong vectơ nhị phân thực vật được liên hợp đầu tiên. Plasmit pNir3K51 (được mô tả trước đó trong sáng chế US số 6420548) được phân cắt bằng HindIII và EcoRI. Đoạn thu được được tách dòng thành trong pCAMBIA2300 (Cambia, Canberra, Úc) được phân cắt bởi cùng enzym để thu được pCAMBIA-Nir3K51.

Trình tự mã hóa cho Hsp70 và Hsp40 được tạo dòng riêng biệt trong plasmit tiếp nhận pCAMBIANir3K51 bằng phương pháp gắn trên cơ sở PCR được Darveau và cộng sự mô tả. (Methods in Neuroscience 26:77-85 (1995)).

Đối với Hsp40, trình tự mã hóa Msj1 (SEQ ID NO: 41) được khuếch đại bằng RT-PCR từ (ecotyp Rangelander) toàn bộ ARN của lá cỏ linh lăng sử dụng các môi Hsp40Luz.1c (SEQ ID NO: 43) và Hsp40Luz-SacI.1272r (SEQ ID NO: 44). Một quá trình khuếch đại thứ hai được thực hiện với các môi Plasto-443c (SEQ ID NO: 5) và Hsp70Ara-SacI.1956r (SEQ ID NO: 45) với sản phẩm thiết kế 660 (SEQ ID NO: 53; Fig.18) làm khuôn. Các sản phẩm PCR sau đó được trộn và sử dụng làm khuôn cho quá trình khuếch đại thứ ba (phản ứng ghép nối) với các môi Plasto-443c (SEQ ID NO: 5) và Hsp40Luz-SacI.1272r (SEQ ID NO: 44). Đoạn thu được được phân cắt bằng HpaI (trong trình tự khởi đầu phiên mã plastocyanin) và được tách dòng thành trong pCAMBIANir3K51, được phân cắt từ trước bằng HpaI (ở trình tự khởi đầu phiên mã Nir) và SacI, và được cắt bằng T4 ADN polymeraza để tạo ra các đầu cùn. Các dòng thu được được sàng lọc để định hướng chính xác và được giải trình tự đối với toàn bộ chuỗi. Plasmit thu được, được đặt tên là R850, được thể hiện trên Fig.83 (SEQ ID NO: 77). Vùng mã hóa của Athsp70-1 được khuếch đại bằng RT-PCR từ ARN của lá Arabidopsis ARN sử dụng các môi Hsp70Ara.1c (SEQ ID NO: 46) và Hsp70Ara-SacI.1956r (SEQ ID NO: 47). Một quá trình khuếch đại thứ hai được thực hiện với các môi Plato-443c (SEQ ID NO: 5) và Hsp70Ara-SacI.1956r (SEQ ID NO: 48) với sản phẩm thiết kế 660 (SEQ ID NO: 53; Fig.18) làm khuôn. Các sản phẩm PCR sau đó được trộn và sử dụng làm khuôn cho quá trình khuếch đại thứ ba (phản ứng ghép nối) với các môi Plasto-443c (SEQ ID NO: 5) và Hsp70Ara-SacI.1956r (SEQ ID NO: 47). Đoạn thu được được phân cắt bằng HpaI (ở trình tự khởi đầu phiên mã plastocyanin) và được tách dòng thành trong pCAMBIANir3K51 được phân cắt bằng HpaI (ở trình tự khởi đầu phiên mã Nir) và SacI và được cắt bằng T4 ADN polymeraza để tạo ra các đầu cùn. Các dòng thu được được

sang lọc để định hướng chính xác và được giải trình tự đối với toàn bộ chuỗi. Plasmid thu được, được đặt tên là R860, được thể hiện trên Fig.84 (SEQ ID NO: 78).

Plasmid biểu hiện Hsp kép được liên hợp như sau. SacI-PlasTer.c (SEQ ID NO: R860) được phân cắt bằng BsrBI (xuôi chiều tác nhân kết thúc NOS), được xử lý bằng T4 ADN polymeraza để tạo ra một đầu bằng, và được phân cắt bằng SbfI (ngược chiều trình tự khởi đầu phiên mã NIR/Plasto khảm). Đoạn thu được (Nir/Plasto trình tự khởi đầu phiên mã thể khảm-trình tự mã hóa HSP70 –nhân tố kết thúc Nos) được tách dòng thành R850 (SEQ ID NO: 77; Fig.42) được phân hủy trước bởi SbfI và SmaI (cả hai được định vị trong nhiều vị trí tách dòng ngược chiều với trình tự khởi đầu phiên mã Nir/Plasto thể khảm). Plasmid thu được, được đặt tên là R870, được thể hiện trên Fig.85 (SEQ ID NO: 79).

Sự liên hợp của các catxet biểu hiện khác

Catxet biểu hiện HcPro

Sản phẩm thiết kế HcPro (35HcPro) được điều chế như Hamilton và cộng sự mô tả. (2002). Tất cả các dòng được sắp xếp theo trình tự để xác định tính nguyên vẹn của các sản phẩm thiết kế. Các plasmid được sử dụng để biến đổi *Agrobacterium tumefaciens* (AGL1; ATCC, Manassas, VA 20108, USA) bằng kỹ thuật xung điện (Mattanovich và cộng sự., 1989). Tính toàn vẹn của tất cả các chủng *A. tumefaciens* được xác nhận bằng phương pháp đánh dấu điểm giới hạn.

Catxet biểu hiện P19

Trình tự mã hóa của protein 19 của virus chùn ngọn tàn lụi cà chua (TBSV) được liên kết với đoạn biểu hiện plastoxyanin ở cổ linh lăng bằng phương pháp gắn trên cơ sở PCR được Darveau và cộng sự mô tả. (Methods in Neuroscience 26: 77-85(1995)). Trong chu trình PCR đầu tiên, một đoạn của trình tự khởi đầu phiên mã plastoxyanin được khuếch đại bằng cách sử dụng các môi Plasto-443c (SEQ ID NO: 5) và Hsp70Ara-SacI.1956r (SEQ ID NO: 49) với sản phẩm thiết kế 660 (SEQ ID NO: 53) làm khuôn. Song song, một đoạn khác chứa trình tự mã hóa p19 được khuếch đại bằng các môi supP19-1c (SEQ ID NO:125) và SupP19-SacI.r (SEQ ID NO: 50) và Hsp70Ara-SacI.1956r (SEQ ID NO: 51) bằng cách sử dụng sản phẩm thiết kế 35S:p19 như được Voinnet và cộng sự mô tả. (The Plant Journal 33: 949-956 (2003)) làm khuôn. Các sản phẩm khuếch đại sau đó được trộn và sử dụng làm khuôn cho chu kỳ khuếch đại thứ hai

(phản ứng ghép nối) với các môi Plasto-443c (SEQ ID NO: 5) và Hsp70Ara-SacI.1956r (SEQ ID NO: 51). Đoạn thu được được phân cắt bằng BamIII (trong trình tự khởi đầu phiên mã plastocyanin) và SacI (ở đầu 3' của đoạn) và được tách dòng thành trong pCAMBIAPlasto được phân cắt từ trước với cùng các enzym. 53; Fig.18), được phân hủy trước bởi các enzym giới hạn giống nhau to give sản phẩm thiết kế số R472. Plasmid R472 được thể hiện trên Fig.86.

Sản phẩm thiết kế số 443

Sản phẩm thiết kế số 443 tương ứng với pCAMBIA2300 (vector trống).

Bảng 1. Các môi oligonucleotit được sử dụng để gắn các catxet biểu hiện

SEQ ID	Tên	Trình tự oligonucleotit (5'---3')
1	XmaI-pPlas.c	AGTTCCTCCGGCTGGTATATTTATATGTTGTC
2	SacI-ATG-pPlas.r	AATAGAGCTCCATTTTCTCTCAAGATGATTAATTAATTAATTAGTC
3	SacI-PlasTer.c	AATAGAGCTCGTTAAATGCTTCTTCGTCTCCTATTTATAATATGG
4	EcoRI-PlasTer.r	TTACGAATTCTCCTTCTAATTGGTGTACTATCATTATCAAGGGGA
5	Plasto-443c	GTATTAGTAATTAGAATTTGGTGTC
6	SpHA(Ind)-Plasto.r	GCAAGAAGAAGCACTATTTTCTCCATTTTCTCTCAAGATGATTA
7	Plasto-SpHA.c	TTAATCATCTTGAGAGAAAATGGAGAAAATAGTGCTTCTTCTTGC
8	HA(Ind)-Sac.r	ACTTTGAGCTCTTAAATGCAAATTCTGCATTGTAACGA
9	pBinPlus.2613c	AGGAAGGGAAGAAAGCGAAAGGAG
10	Mut-ATG115.r	GTGCCGAAGCACGATCTGACAACGTTGAAGATCGCTCACGCAAGAAAG ACAAGAGA
11	Mut-ATG161.c	GTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAG CGA
12	LC-C5-1.110r	TCTCCTGGAGTCACAGACAGGGTGG
13	E1 H1B-E1 H5L.r	TCATAGTCAGCGAAATGCCCTGGGTAACAGAGGTCATTGGTTGGATTG GCCT
14	E1 H5N-E1 H1B.c	ATGACCTCTGTTACCCAGGGCATTTCGCTGACTATGAGGAAGTGAAGG
15	E2 H5I-RB H1B.r	CCAATTCACTTTTTATAATTCTGATCCAAAGCCTCTACTCAGTGCGA
16	RB H1B-E2 H5L.c	GGCTTTGGATCAGGAATTATGAAAAGTGAATTGGAATATGGTAACTGC AAC
17	E1 H1B-F'1 H5L.r	GGCTATTCCTTTTAATAGGCAGAGCTTCCCGTTGTGTGTCTTTTCCAG
18	F'1 H5N-E1 H1B.c	AACGGGAAGCTCTGCCTATTAAGGAATAGCCCCACTACAATTGGGT
19	F'2 H5I-E2 H1B.r	GGAGTTTGACACTTGGTGTGCAATTTATCCATTGGTGCATTTGAGTTG
20	E2 H1B-F'2 H5L.c	AATGCACCAATGGATAAATGCAACACCAAGTGTCAACTCCAATGGGG

SEQ ID	Tên	Trình tự oligonucleotit (5'---3')
21	E1 H5I-E1 H1NC.r	TCTTCATAGTCGTTGAAACTCCCTGGGTAACATGTTCCATTCTCAGGA
22	E1 H1NC-E1 H5I.c	CTGAGAATGGAACATGTTACCCAGGGAGTTTCAACGACTATGAAGAAC
23	E2 H1NC-RB H5I.r	ATTTGAGGTGATGATTGCTGAGTCCCCTTTCTTGACAATTTTGTATGCA TA
24	RB H5I-E2 H1NC.c	GTCAAGAAAGGGGACTCAGCAATCATCACCTCAAATGCACCAATGGAT
25	HA-SacI.r	TTAACTTAGAGCTCTTAGATGCATATTCTACACTGCAAAGACC
26	ApaI-H1B.c	TGTCGGGGCCCATGAAAGTAAACTACTGGTCCTGTTATGCACATT
27	StuI-H1B.r	AAATAGGCCTTTAGATGCATATTCTACACTGTAAAGACCCATTGGA
28	SpPDI-H1B.c	TTCTCAGATCTTCGCTGACACAATATGTATAGGCTACCATGCTAACAAC
29	SacI-H1B.r	CTTAGAGCTCTTAGATGCATATTCTACACTGTAAAGACCCATTGGAA
30	ApaI-SpPDI.c	TTGTCGGGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGT
31	ApaI-H5 (A-Indo).1c	TGTCGGGGCCCATGGAGAAAATAGTGCTTCTTCTTGCAAT
32	H5 (A-Indo)-StuI.1707r	AAATAGGCCTTTAAATGCAAATTCTGCATTGTAACGA
33	H3B-SpPDI.r	TGTCATTTCCGGGAAGTTTTTGAGCGAAGATCTGAGAAGGAACCA
34	SpPDI-H3B.c	TCTCAGATCTTCGCTCAAAAACCTCCCGGAAATGACAACAGCACG
35	StuI-H3B.r	AAAATAGGCCTTCAAATGCAAATGTTGCACCTAATGTTGCCTTT
36	TmD H5I-H3B.r	ATTTGGTAAGTTCTATTGACTTCAGCTCAACGCCCTTGATCTGG
37	H3B-TmD H5I.c	TGAGCTGAAGTCAATAGGAACTTACCAAATACTGTCAATTTATTCAAC
38	HBf-SpPDI.r	GTTATTCCAGTGCAGATTTCGATCAGCGAAGATCTGAGAAGGAACCAAC AC
39	SpPDI-HBf.c	CAGATCTTCGCTGATCGAATCTGCACTGGAATAACATCTTCAAACCTCAC C
40	StuI-HBf.r	AAAATAGGCCTTTATAGACAGATGGAGCATGAAACGTTGTCTCTGG
41	TmD H5I-B Flo.r	TGACAGTATTTGGTAGTTATCCAATCCATCATCATTTAAAGATGC
42	B Flo-TmD H5I.c	GGATTGGATAACTACCAAATACTGTCAATTTATTCAACAGTGGCGAGTT C
43	Hsp-40Luz.1c	ATGTTTGGGCGCGGACCAAC
44	Hsp40Luz-SacI.1272r	AGCTGAGCTCCTACTGTTGAGCGCATTGCAC
45	Hsp40Luz-Plasto.r	GTTGGTCCGCGCCCAAACATTTTCTCTCAAGATGAT
46	Hsp70Ara.1c	ATGTCGGGTAAAGGAGAAGGA
47	Hsp70Ara-SacI.1956r	AGCTGAGCTCTTAGTCGACCTCCTCGATCTTAG
48	Hsp70Ara-Plasto.r	TCCTTCTCCTTTACCCGACATTTTCTCTCAAGATGAT
49	supP19-plasto.r	CCTTGTATAGCTCGTTCCATTTTCTCTCAAGATG
50	supP19-1c	ATGGAACGAGCTATACAAGG
51	SupP19-SacI.r	AGTCGAGCTCTTACTCGCTTTCTTTTTCGAAG

SEQ ID	Tên	Trình tự oligonucleotit (5'---3')
89	PacI-MCS-2X35S.c	AATTGTTAATTAAGTCGACAAGCTTGCATGCCTGCAGGTCAAC
90	CPMV 5'UTR- 2X35S.r	TCAAAACCTATTAAGATTTTAATA CCTCTCCAAATGAAATGAACTTCC
91	2X35S -CPMV 5'UTR.c	TTGGAGAGGT TATTAATAATCTTAATAGGTTTTGATAAAAGCGAACGTGGG
92	ApaI -M prot.r	TCTCCAT GGGCCCC GACAAATTTGGGCAGAATATACAGAAGCTTA

Bảng 2: Đoạn biểu hiện được sử dụng để biểu hiện các chất ngưng kết hồng cầu cúm với các peptit tín hiệu gốc hoặc PDI.

Chủng Agro	HA được biểu hiện	Peptit tín hiệu	Catxet biểu hiện
AGL1/540	H1 (A/New Caledonia/20/99)	PDI	Plastoxyanin
AGL1/774	H1 từ A/Brisbane/59/2007	gốc	Plastoxyanin
AGL1/787	H1 từ A/Brisbane/59/2007	PDI	Plastoxyanin
AGL1/732	H1 (A/Brisbane/59/2007)	gốc	35S/CPMV-HT
AGL1/736	H3 (A/Brisbane/10/2007)	PDI	35S/CPMV-HT
AGL1/660	Tinh sạch H5 VLP (A/Indonesia/5/2005)	gốc	Plastoxyanin
AGL1/739	B (B/Florida/4/2006)	PDI	35S/CPMV-HT
AGLI/828	CPMV HT-LC C51	C51LC	35S/CPMV-HT
AGLI/690	H1/Bris RB + H5/Indo SDC	gốc	Plasto
AGLI/691	H1/Bri E1-RB-E2 + H5 SDC	gốc	Plasto
AGLI/696	H5/Indo RB+H1/NC SDC	PDI	Plasto
AGLI/733	H1/Bri	PDI	35S/CPMV-HT
AGLI/734	H1/Bri RB + H5/Indo SDC	gốc	35S/CPMV-HT
AGLI/737	Vùng ngoài H3/Bri +H5/Indo TDC	PDI	35S/CPMV-HT
AGLI/745	Vùng ngoài B/Flo_ H5/Indo TDC	PDI	35S/CPMV-HT
AGL1/747	Vùng ngoài B/Flo_ H5/Indo TDC	PDI	2X35S/CPMV-HT

Sản xuất sinh khối của cây, chất cấy truyền, thâm nhiễm agro, và thu hoạch

Cây *Nicotiana benthamiana* được trồng từ hạt trong các ngăn nông được đổ đầy chất nền rêu than bùn thương mại. Các cây được trồng trong nhà kính dưới chu kỳ sáng 16/8 và chế độ nhiệt độ là 25°C ngày/20°C đêm. Ba tuần sau khi gieo hạt, nhổ các cây non riêng lẻ ra, trồng lại trong các chậu và để phát triển trong nhà kính trong khoảng thời gian ba tuần nữa dưới các điều kiện môi trường như nhau. Trước khi biến nạp, các chồi đỉnh và

chồi nách được loại ra tại các thời điểm khác nhau như được chỉ ra dưới đây, hoặc bằng cách ngắt các chồi từ cây, hoặc bằng cách xử lý hóa học cây.

Agrobacteria được gây nhiễm bằng mỗi sản phẩm thiết kế được phát triển trong môi trường YEB có bổ sung axit 2-[N-morpholino]etansulfonic (MES) 10mM, axetosyringon 20 μ M, kanamycin 50 μ g/ml và carbenixilin 25 μ g/ml pH=5,6 cho đến khi chúng đạt đến OD₆₀₀ nằm trong khoảng 0,6 và 1,6. Huyền phù *Agrobacterium* được ly tâm trước khi sử dụng và được tạo huyền phù lại trong môi trường thâm nhiễm (MgCl₂ 10mM và MES 10mM pH=5,6). Sự thâm thấu qua xi lanh được tiến hành như được mô tả bởi Liu và Lomonosoff (2002, *Journal of Virological Methods*, 105:343-348). Nhờ thâm nhiễm bằng chân không, huyền phù *A. tumefaciens* được ly tâm, tạo huyền phù lại trong môi trường thâm nhiễm và được dự trữ qua đêm ở nhiệt độ 4°C. Vào ngày thâm nhiễm, các mẻ canh trường được làm loãng trong 2,5 thể tích canh trường và được làm ấm trước khi sử dụng. Toàn bộ các cây *N. benthamiana* hoặc *N. tabacum* were được đặt lộn ngược trong huyền phù vi khuẩn trong thùng thép không gỉ kín không khí dưới chân không có 2666,4-5332,9Pa trong khoảng thời gian 2-phút. Sự thâm nhiễm chân không hoặc ống tiêm, các cây được đưa trở lại nhà kính trong khoảng thời gian từ 4 đến 5 ngày trong chu kỳ ủ cho đến khi thu hoạch. Trừ khi được quy định theo cách khác, tất cả các quá trình ủ ấm được thực hiện là quá trình đồng thâm thấu với AGL1/35S-HcPro theo tỷ lệ 1:1, ngoại trừ đối với các chủng mang catxet CPMV-HT mà được đồng thâm nhiễm với chủng AGL1/R472 theo tỷ lệ 1:1.

Lấy mẫu lá và tổng dịch chiết protein

Sau khi ủ, phần khí sinh của cây được thu lại, đông lạnh ở nhiệt độ -80°C, nghiền thành các mẫu. Tổng protein hòa tan được chiết xuất bằng cách làm đồng nhất (Polytron) mỗi mẫu nguyên liệu thực vật đã được làm đông lạnh- được nghiền trong 3 thể tích chứa Tris 50mM pH=8, NaCl 0,15M, và phenylmetansulfonyl florua 1mM. Sau khi đồng nhất hóa, huyền phù đặc được ly tâm với 20000g trong khoảng thời gian 20 phút ở nhiệt độ 4°C và dịch chiết thô được lọc này (phần nổi trên mặt) được giữ để phân tích. Hàm lượng protein tổng số của dịch chiết thô đã lọc được xác định bằng thử nghiệm Bradford (Bio-Rad, Hercules, CA) sử dụng albumin huyết thanh bò làm mẫu chuẩn.

Phân tích protein và miễn dịch thấm

Nồng độ protein được xác định bằng phép phân tích protein BCA (Pierce Biochemicals, Rockport IL). Các protein được tách ra bằng SDS-PAGE dưới các điều kiện khử và được nhuộm với Coomassie Blue. Gel đã nhuộm được quét và phép phân tích đo tỷ trọng được tiến hành bằng cách sử dụng phần mềm ImageJ (NIH).

Các protein từ phân đoạn rửa giải từ SEC được kết tủa bằng axeton (Bollag và cộng sự, 1996), được tái tạo huyền phù trong 1/5 thể tích đệm cân bằng/rửa giải và được phân tách bằng SDS-PAGE dưới các điều kiện khử và được chuyển electron lên các màng polyvinyliden diflorua (PVDF) (Roche Diagnostics Corporation, Indianapolis, IN) để phát hiện miễn dịch. Trước khi thấm miễn dịch, các màng được chặn bởi sữa không kem 5% và Tween-20 0,1% trong nước muối đệm Tris (TBS-T) trong khoảng thời gian từ 16 đến 18 giờ ở nhiệt độ 4°C.

Miễn dịch thấm được thực hiện bằng cách ủ với một kháng thể thích hợp (Bảng 6), ở 2µg/ml trong sữa phân lập 2% trong TBS-Tween 20 0,1%. Các kháng thể thứ phát được sử dụng cho sự phát hiện bằng phát quang hóa học như được thể hiện trong bảng 4, pha loãng như được thể hiện trong sữa phân lập 2% trong TBS-Tween 20 0,1%. Các phức chất phản ứng miễn dịch được phát hiện nhờ sự phát quang hóa học sử dụng luminol làm chất nền (Roche Diagnostics Corporation). Enzym peroxidaza ở cây cải ngựa liên hợp kháng thể IgG ở người được thực hiện bằng cách sử dụng kit liên hợp Peroxidaza được hoạt hóa EZ-Link Plus® (Pierce, Rockford, IL). Virut bị bất hoạt, toàn diện (Toàn bộ, virut được bất hoạt-WIV), được sử dụng làm các đối chứng phát hiện các typ phụ H1, H3 và B, được mua từ National Institute for Biological Standards and Control (NIBSC).

Bảng 3: Các điều kiện điện chuyển, các kháng thể, và các sự pha loãng để miễn dịch thấm các protein biểu hiện

Typ phụ HA	Chủng cúm	Điều kiện điện di	Kháng thể ban đầu	Độ pha loãng	Kháng thể thứ hai	Độ pha loãng
H1	A/Brisbane/59/2007 (H1N1)	Khử	FII 10-I50	4µg/ml	Chuột kháng dê (JIR 115-035-146)	1:10 000
H1	A/New Caledonia/20/99 (H1N1)	Khử	FII 10-I50	4µg/ml	Chuột kháng dê (JIR 115-035-146)	1:10 000
H3	A/Brisbane/10/2007 (H3N2)	Không khử	NIBSC 08/124	1:594000.	Cừu kháng thỏ (JIR 313-035-045)	1:10 000

H5	A/Indonesia/5/2005 (H5N1)	Khử	ITC IT-003-005V	1:59400 0.	Thỏ kháng dê (JIR 111-035-144)	1:10 000
B	B/Florida/4/2006	Không khử	NIBSC 07/356	1:2000	Cừu kháng thỏ (JIR 313-035-045)	1:10 000

FII: Fitzgerald Industries International, Concord, MA, USA;
 NIBSC: National Institute for Biological Standards and Control;
 JIR: Jackson ImmunoResearch, West Grove, PA, USA;
 ITC: Immune Technology Corporation, Woodside, NY, USA;

Lọc trong và cô đặc trước khi SEC

Để cải thiện sự phân giải và làm tăng tín hiệu trong các phần rửa giải, các phần chiết được nạp trên phương pháp sắc ký rây phân tử, dịch chiết protein thô được lọc và cô đặc sử dụng phương pháp dưới đây. Dịch chiết được ly tâm tại 70 000g, nhiệt độ 4°C trong khoảng thời gian 20 phút và pelet được rửa hai lần bằng cách tạo huyền phù lại trong 1 thể tích (được so sánh với thể tích dịch chiết ban đầu) của đệm chiết (Tris 50mM pH=8, NaCl 0,15M) và ly tâm tại 70 000g, nhiệt độ 4°C trong khoảng thời gian 20 phút. Pelet thu được được tạo huyền phù lại trong 1/3 thể tích (được so sánh với thể tích chiết ban đầu) và các protein (gồm các VLP) được kết tủa bằng cách bổ sung PEG 3350 20% (trọng lượng/ thể tích) sau đó bằng cách ủ trên đá trong khoảng thời gian 1 giờ. Các protein kết tủa được tái sinh lại bằng ly tâm tại 10 000g, 4°C, 20 phút, và được tạo huyền phù lại trong 1/15 thể tích (được so sánh với thể tích chiết ban đầu) của đệm chiết. Sau khi hoàn thiện sự tạo huyền phù lại protein, tiến hành ly tâm lần cuối tại 20 000g, 4°C, 5 phút tạo pelet không hòa tan và thu hồi phần nổi trên mặt sạch.

Sắc ký loại trừ kích thước dịch chiết protein

Các cột sắc ký loại trừ kích thước (Phương pháp sắc ký rây phân tử-SEC) chứa các hạt có độ phân giải cao 32 ml Sephacryl™ S-500 (S-500 HR: GE Healthcare, Uppsala, Sweden, Cat. GE Healthcare, Uppsala, Sweden, Cat.Số 17-0613-10) được tập hợp và làm cân bằng bằng đệm cân bằng/rửa giải (Tris 50mM pH=8, NaCl 150mM). 1,5mililit dịch chiết protein thô được nạp lên trên cột tiếp theo bước rửa giải bằng 45mL đệm cân bằng/rửa giải. Các dịch rửa giải này được thu trong các đoạn có 1,5mL hàm lượng protein tương đối của các đoạn rửa giải được giám sát bằng cách trộn 10µL đoạn với 200µL thuốc thử nhuộm màu protein Bio-Rad được pha loãng (Bio-Rad, Hercules, CA cột được rửa bằng 2 thể tích cột NaOH 0,2N tiếp theo là 10 thể tích cột Tris pH8 50mM, NaCl 150mM, etanol 20%. Mỗi quá trình phân tách được theo sau bởi một quá trình hiệu chuẩn cột bằng

Blue Dextran 2000 (GE Healthcare Bio-Science Corp., Piscataway, NJ, USA). Các dữ liệu rửa giải của Blue Dextran 2000 và các protein hòa tan chình được so sánh giữa mỗi quá trình phân tách để đảm bảo độ đồng nhất của các dữ liệu rửa giải giữa các cột được sử dụng.

Ví dụ thực hiện sáng chế

Ví dụ 1: Chiến lược hoán đổi vùng cho các vùng RB và/hoặc Esteraza trong các tế bào gốc typ phụ cúm.

Vùng phụ RB của H5/Indo có thể được thay thế bởi vùng phụ RB của H1, H3 hoặc B HA. HA thể khảm thu được tạo ra SDC H5/Indo tạo thành các VLP và thể hiện vùng phụ RB bao gồm các vị trí tạo miễn dịch của H1, H3 hoặc B. Vùng phụ H5/Indo RB có thể được chèn trên tế bào gốc H1 (H1/NC). Các Fig.15A và 15B minh họa các trình tự axit amin tại các vị trí dung hợp các vùng phụ được chỉ ra, và các trình tự axit amin của các vùng phụ điển hình được tạo ra trên Fig.2 (sản phẩm thiết kế 690, 734, 696 và 691) và các Bảng 4 (sản phẩm thiết kế 900 và 745) và 5 (sản phẩm thiết kế 910, 920 và 930). Các trình tự axit amin được minh họa trên Fig.2, và các Bảng 4 và 5 không gồm các trình tự peptit tín hiệu.

Bảng 4 Các vùng phụ và HA cúm thể khảm. HA cúm thể khảm bao gồm vùng phụ RB khác loại.

Sản phẩm thiết kế số	Tế bào gốc đầu tận cùng N: -F'1+E1	Đầu RB	Tế bào gốc đầu tận cùng C: E2+F'2-Kết thúc
900 SEQ ID NO: 105)	H5/Indo DQICIGYHA NNSTEQVDTI MEKNVTVTTH AQDILEKTH NGKLCDLDG VKPLILRDCS VAGWLLGNP MCDEFINVE WSYIVEKAN PTNDLCYP	H3/Brisbane YDVPDYASLRSLVASSGTLEFN NESFNWTGVTQNGTSSACIRRS NNSFFSRLNWLTHLKFKYPALN VTMPNNEKFDKLYIWGVHHPG TDNDQIFLYAQASGRITVSTKRS QQTVIPNIGSRPRVRNIPSRSIY WTIVKPGDILLINSTGNLIAPRG Y FKIRSGKSS	H5/Indo IMKSELEYGNCNTKCQTPMGA INSSMPFHNHPLTIGECPKY VKS NRLVLATGLRNSPQRESR RKKRGLFGAIAAGFIEGGWQG MVDGWYGYHHSNEQSGSYA ADKESTQKAIDGVTNKVNSII DKMNTQFEAVGREFNNLERRI ENLNKKMEDGFLDVWTYNA ELLVLMENERTLDFHDSNVKN LYDKVRLQLRDNAKELGNG CFEFYHKCDNECMESIRNGTY NYPQYSEEARLKREEISGVKLE SIGTYQILSIYSTVASSLALAIM MAGLSLWMCSNGSLQCRICI
745	H5/Indo	B/Florida	H5/Indo

(SEQ ID NO: 106)	DQICIGYHA NNSTEQVDI MEKNVTVTH AQDILEKTH NGKLCDLDG VKPLILRDCS VAGWLLGNP MCDEFINVPE WSYIVEKAN PTNDLCYP	IMHDRTKIRQLPNLLRGYENIRL STQNVIDAEKAPGGPYRLGTSG SCPNATSKSGFFATMAWAVPK DNNKNATNPLTVEVPYICTEGE DQITVWGFHSDNKTQMKNLYG DSNPQKFTSSANGVTTHYVSQI GSFPDQTEDGGLPQSGRIVVDY MMQKPGKTGTIVYQRGVLLPQ KWCASGRSK	IMKSELEYGNCNTKCQTPMGA INSSMPFHNIHPLTIGECPKYV KSNRLVLATGLRNSPQRESRRK KRGLFGAIAGFIEGGWQGMV DGWYGYHHSNEQSGGYAADK ESTQKAIDGVTNKNVNSIIDKM NTQFEAVGREFNNLERRIENLN KKMEDGFLDVWTYNAELLV LMENERTLDFHDSNVKNLYDK VRLQLRDNAKELGNGCFEY HKCDNECMESIRNGTYNYPQY SEEARLKREEISGVKLESIGTY QILSIYSTVASSLALAIMMAGLS LWMCSNGSLQCRICI
-------------------------	---	---	---

Axit amin 1-92 có SEQ ID NO: 105 là vùng F'1+E1 của H5/Indo; các axit amin 93-259 là vùng đầu RB của H3/Brisbane; các axit amin 260-548 là vùng E2+F'2 của H5/Indo.

Axit amin 1-92 có SEQ ID NO: 106 là vùng F'1+E1 của H5/Indo; các axit amin 93-276 là vùng đầu RB của B/Florida; các axit amin 277-565 là vùng E2+F'2 của H5/Indo.

Bảng 5 Các vùng phụ và HA cúm thể khảm. HA cúm thể khảm bao gồm vùng phụ RB khác loại.

Sản phẩm thiết kế số	Tế bào gốc đầu tận cùng N: F'1	Đầu E1-RB-E2	Tế bào gốc đầu tận cùng C: F'2- Kết thúc
910 (SEQ ID NO: 107)	H5/Indo DQICIGYHA NNSTEQVD TIMEKNVT VTHAQDILE KTHNGKLC	H3/Brisbane DSPHQILDGENCTLIDA LLGDPQCDGFQNKKW DLFVERSKAYSNCYPY DVPDYASLRSLVASSG TLEFNNEFNTWTGVTQ NGTSSACIRRSNNSFFS RLNWLTHLKFKYPALN VTMPNNEKFDKLYIWG VHHPGTDNDQIFLYAQ ASGRITVSTKRSQQTVI PNIGSRPRVRNIPSRISIY WTIVKPGDILLINSTGN LIAPRGYFKIRSGKSSIM RSDAPIGK	H5/Indo CNTKCQTPMGAINSSMPFHNIHPLTIGEC CPKYVKSNRLV LATGLRNSPQRESRRKKRGLFGAIAGF IEGGWQGMVD GWYGYHHSNEQSGGYAADKESTQKA IDGVTNKNVNSII DKMNTQFEAVGREFNNLERRIENLNK KMEDGFLDVW TYNAELLVLMENERTLDFHDSNVKNLY YDKVRLQLRDN AKELGNGCFEYHKCDNECMESIRNG TYNYPQYSEEA RLKREEISGVKLESIGTYQILSIYSTVAS SLALAIMMAGL SLWMCSNGSLQCRICI

920 (SEQ ID NO: 108)	H5/Indo DQICIGYHA NNSTEQVD TIMEKNVT VTHAQDILE KTHNGKLC	B/Florida PDCLNCTDL DVALGRP MCVGTTPSAKASILHE VKPVTSGCFPIMHDRT KIRQLPNLLRGYENIRL STQNVIDAEKAPGGPY RLGTSGSCPNATSKSGF FATMAWAVPKDNNKN ATNPLTVEVPYICTEGE DQITVWGFHSDNKTQM KNLYGDSNPQKFTSSA NGVTTHYVSQIGSFPDQ TEDGGLPQSGRIVVDY MMQKPGKTGTIVYQR GVLLPQKVWCASGRSK VIKGSPLIGEAD	H5/Indo CQTPMGAINSSMPFHNIHPLTIGECPKY VKS NRLVLATG LRNSPQRESRRKKRGLFGAIAAGFIEGG WQGMVDGWYG YHHSNEQSGSGYAADKESTQKAIDGVT NKVNSIIDKMN TQFEAVGREFNNLERRIENLNKKMED GFLDVWTYNAE LLVLMENERTLDFHDSNVKNLYDKVR LQLRDNAKELG NGCFEFYHKCDNECMESIRNGTYNYP QYSEEARLKREE ISGVKLESIGTYQILSIYSTVASSLALAI MMAGLSLWMC SNGSLQCRICI
930 (SEQ ID NO: 109)	H1/NC DTICIGYHA NNSTDTVD TVLEKNVT VTHSVNLL EDSHNGKL C	H5/Indo DLDGVKPLILRDCSVA GWLLGNPMCDEFINVP EWSYIVEKANPTNDLC YPGSFNDYEELKHLLSR INHFEKIQIIPKSSWSDH EASSGVSSACPYLGSPS FFRN VVWLIIKKNSTYP TIKKSYNNTNQEDLLV LWGIHHPNDAAEQTRL YQNPTTYISIGTSTLNQ RLVPKIATR SKVNGQS GRMEFFWTILKPNDAIN FESNGNFIAPEYAYKIV KKGDSAIMKSELEYGN	H1/NC CDAKCQTPQGAINSSLPFQNVHPVTIG ECPKYVRS AKL RMVTGLRNIPSIQSRGLFGAIAAGFIEGG WTGMVDGWYG YHHQNEQSGSGYAADQKSTQNAINGIT NKVNSVIEKMN TQFTAVGKEFNKLERRMENLNKKVDD GFLDIWTYNAE LLVLENER TLDFHDSNVKNLYEKVK SQLKNNAKEIGN GCFEFYHKCNNECMESVKNGTYDYPK YSEESKLNREK IDGVKLESMGVYQILAIYSTVASSLVL LVSLGAISFWMC SNGSLQCRICI

Axit amin 1-42 có SEQ ID NO: 107 là vùng F'1+E1 của H5/Indo; các axit amin 43-228 là vùng đầu E1-RB-E2 của H1/Brisbane và axit amin 229-507 là vùng F'2 của H5/Indo.

Axit amin 1-42 có SEQ ID NO: 108 là vùng F'1+E1 của H5/Indo; các axit amin 43-281 là vùng đầu E1-RB-E2 của H1/Brisbane và axit amin 282-556 là vùng F'2 của H5/Indo.

Axit amin 1-42 có SEQ ID NO: 109 là vùng F'1+E1 của H5/NC; các axit amin 43-273 là vùng đầu RB của H5/Indo và axit amin 274-548 là vùng E2+F'2 của H1/NC.

Các điểm dung hợp cho các thể khảm khác nhau được chọn gần với (nhưng không cần thiết trực tiếp tại) các đầu tận cùng N và C của các vùng phụ khác nhau – mà không bị ràng buộc bởi lý thuyết, các điểm dung hợp này được chọn để tối đa độ ổn định của HA thể khảm. Ví dụ, sự bảo tồn cấu trúc và trình tự được quan sát tại đầu tận cùng N của vùng phụ RB (Ha et al. 2002, EMBO J. 21:865-875; được kết hợp ở đây để tham khảo). Vùng biến thiên ít hơn trong trình tự ban đầu được tìm thấy tại bộ ba C-F/Y-P được định vị tại khoảng 15 axit amin trước khi, trong vùng phụ E1. Xystein này được gồm trong cầu disulfua #3, được bảo tồn trong số các HA (tham khảo các Fig.46 và 47). Chỗ nối tại Cys này tạo ra độ ổn định thích hợp, hoặc tốt hơn với HA thể khảm tương ứng với trình tự gốc. Đầu tận cùng C của RB tạo ra các đặc tính bảo tồn: ví dụ, chuỗi tận Ser bảo toàn tại vị trí -1 và vùng phụ E2 bắt đầu với tám beta được bảo tồn trong tất cả HA dựa trên sự liên kết (Ha et al. 2002, EMBO J. 21:865-875; được kết hợp ở đây để tham khảo). Do đó, đầu tận cùng C của RB có thể được dung hợp với axit amin ban đầu có cấu trúc tám beta này của vùng phụ E2. Hơn nữa, mô hình cầu nối disulfua không bị thay đổi, hoặc bị thay đổi lớn, với các thể khảm bao gồm các vùng phụ RB của H1/NC, H1Bri, H3/Bri, hoặc B/Flo dựa trên H5/Indo SDC, và với vùng phụ H5/Indo RB dựa trên H1 SDC (tổng số bằng 6), nhưng cầu nối disulfua sẽ được bổ sung (cầu #8) trên HA lai của B RB trên tế bào gốc H5. Sự bổ sung cầu nối disulfua này không gây cản trở đến sự gấp nếp của HA (do nó được định vị trong vùng RB và Cys liền kề trên trình tự), và có thể được tạo ra ngay cả với HA lai ổn định hơn.

Các vùng phụ E1-RB-E2 thuộc typ HA cúm thứ nhất được thay thế bởi các vùng phụ E1-RB-E2 thuộc typ cúm thứ hai. Sự sắp xếp này thể hiện số lượng axit amin lớn hơn thuộc typ thứ hai tại bề mặt của H5-VLP. Trong ví dụ này, HDC của H1, H3 hoặc B được đặt trên H5/Indo SDC, và HDC của H5/Indo trên H1/NC SDC (Bảng 5).

Mỗi nối của HDC được định rõ với chuỗi xystein ở đầu được bảo tồn (bao gồm cầu nối disulfua #6 thuộc HA typ A và #7 trong HA typ B). Mỗi nối của HDC tại đầu tận cùng C của vùng phụ E2 được định rõ bởi chuỗi tận xystein được bảo tồn khác bao gồm cầu nối disulfua #6 (axit amin thứ hai của vùng phụ F'2) thuộc bệnh cúm typ H1 hoặc H3 trên SDC của H5/Indo hoặc với bệnh cúm typ H5 trên SDC của H1. Với bệnh cúm B thể khảm, sự nối được thiết lập mối liên hệ tại Xystein thứ nhất bao gồm cầu nối disulfua #4 (4 axit amin được định vị trên vùng phụ F'2, và được bảo tồn trong số các HA). Các thể

khả thu được không thể hiện sự biến đổi bất kỳ trong các mô hình cầu nối disulfua - H1/H3/H5 các HA lai sẽ chứa 6 cầu nối disulfua và B lai sẽ có 7 trong số chúng.

Ví dụ 2: Sự thay thế các vùng phụ liên kết thụ thể (RB) hoặc liên kết thụ thể và esteraza (E1-RB-E2) của H5 A/Indonesia/5/05 với vùng phụ của H1 A/Brisbane/59/2007: So sánh sự biểu hiện của các dạng gốc và dạng thể khả.

Để liên kết mức tích tụ cao của các VLP từ H5 A/Indonesia/5/05 với các đặc tính có tính kháng nguyên của H1 A/Brisbane/59/2007, chất ngưng kết hồng cầu thể khả được thiết kế bao gồm các vùng từ H1 A/Brisbane/59/2007 được dung hợp với đám vùng gốc H5 A/Indonesia/5/05. Các catxet biểu hiện để hiểu hiện các điểm dung hợp H5/H1 chất ngưng kết hồng cầu được thể hiện trên Fig.1 và trình tự axit amin của các protein dung hợp hoàn chỉnh được tạo ra được minh họa trên Fig.2.

Để so sánh với mức tích tụ của các chất ngưng kết hồng cầu thể khả H5/H1 với chất ngưng kết hồng cầu thuộc dạng gốc của chúng, các cây *Nicotiana benthamiana* được thâm nhiễm bởi AGL1/774, AGL1/691 và AGL1/690, và các lá được thu sau khoảng thời gian ủ sáu ngày. Để xác định mức kết thụ của mỗi dạng HA trong các lá được thâm nhiễm vi khuẩn Agro, các protein được chiết từ mô lá được thâm nhiễm và được phân tích bằng thẩm tách Western sử dụng các kháng thể đơn dòng kháng HA. Băng duy nhất có khoảng 75 kDa (Fig.3), tương ứng với kích thước có dạng HA0 không phân cắt của chất ngưng kết hồng cầu cúm, được phát hiện trong các dịch chiết từ các lá được thâm nhiễm với AGL1/690 nhưng không phải trong AGL1/774 hoặc AGL1/691, chỉ ra rằng chất ngưng kết hồng cầu thể khả bao gồm vùng liên kết thụ thể của H1 A/Brisbane/59/2007 được dung hợp với đoạn chính H5 A/Indonesia/5/05 được tích tụ với mức cao hơn cả hai dạng gốc của H1 A/Brisbane/59/2007 (AGL1/774) và chất ngưng kết hồng cầu thể khả bao gồm các vùng esteraza và liên kết thụ thể của H1 A/Brisbane/59/2007 với đoạn chính H5 A/Indonesia/5/05. Toàn bộ virus được bất hoạt (WIV) (H1 A/Brisbane/59/2007) được sử dụng làm đối chứng dương được xác định dưới dạng nhiều băng với băng chính tại khoảng 80 kDa, tương ứng với trọng lượng phân tử của tiền chất HA0 của H1 A/Brisbane/59/2007. Các kết quả này chứng minh rằng sự thay thế vùng liên kết thụ thể từ H5 A/Indonesia/5/05 với vùng H1 A/Brisbane/59/2007 được tạo ra chất ngưng kết hồng cầu thể khả được thể hiện vùng kháng nguyên của H1 và được tích tụ với mức cao hơn H1 A/Brisbane/59/2007 gốc trong cây. Tuy nhiên, chất ngưng kết hồng cầu thể khả

trong đó esteraza và vùng liên kết thụ thể từ H5 A/Indonesia/5/05 được thay thế bởi các vùng của H1 A/Brisbane/59/2007 không được tích tụ với mức có thể phát hiện trong cây.

Sự dung hợp vùng liên kết thụ thể từ H1 A/Brisbane/59/2007 đến đoạn chính H5 A/Indonesia/5/05 như phương pháp tích tụ tăng của các VLP trình diện kháng nguyên H1 trong cây được đánh giá lại dưới sự kiểm soát catxet biểu hiện dựa trên CPMV-HT mạnh. Chiến dịch dung hợp cũng được so sánh với sự thay thế peptit tín hiệu sự thay thế có nghĩa là mức tích tụ tăng. Catxet biểu hiện để biểu hiện các điểm dung hợp chất ngưng kết hồng cầu H5/H1 nhờ CPMV-HT được thể hiện trên Fig.8 và trình tự axit amin của protein dung hợp hoàn chỉnh đã tạo ra được thể hiện trên Fig.2.

Cây *Nicotiana benthamiana* được thâm nhiễm bởi AGL1/732, AGL1/733 hoặc AGL1/734, và các lá được thu sau khoảng thời gian ủ sáu ngày. Để xác định mức của mỗi dạng HA trong các lá thâm nhiễm vi khuẩn Agro, protein đầu tiên được chiết từ mô lá được thâm nhiễm và được phân tích bởi kỹ thuật thẩm tách miễn dịch sử dụng các kháng thể đa dòng kháng H1 (Brisbane). Băng duy nhất khoảng 75 kDa (Fig.6), tương ứng với kích thước của dạng HA0 được phân cắt của chất ngưng kết hồng cầu cúm, được phát hiện trong dịch chiết từ lá được thâm nhiễm bởi AGL1/732, AGL1/733 và AGL1/734. Tuy nhiên, mặc dù chất ngưng kết hồng cầu được phát hiện trong tất cả các dịch chiết được phân tích, nên sự khác nhau chủ yếu về sự kết thụ không được lưu ý. Trong khi sự biểu hiện H1 A/Brisbane/59/2007 được phát hiện vừa đủ dưới các điều kiện này khi sử dụng peptit tín hiệu gốc của nó (732), sự thay thế peptit tín hiệu với peptit PDI được kết quả trong sự tích tụ cao hơn của H1 A/Brisbane/59/2007 gốc (733), và chất ngưng kết hồng cầu H5/H1 thể khảm (734) được tích tụ với mức cao nhất. Nói chung, các kết quả này chỉ ra rằng sự dung hợp vùng liên kết thụ thể từ H1 trên đoạn chính H5 dẫn đến sự tích tụ cao của chất ngưng kết hồng cầu trình diện kháng nguyên H1 và mức tích tụ thu được với sự dung hợp trong các cây này cao hơn thu được với dạng gốc với sự thay thế hoặc không có sự thay thế peptit tín hiệu.

Ví dụ 3: Sự thay thế vùng phụ liên kết thụ thể (RB) của H1 A/New Caledonia/20/99 với H5 A/Indonesia/5/05. So sánh với sự biểu hiện của các dạng gốc và thể khảm.

Sử dụng đoạn chính H1 (từ A/New Caledonia/20/99) để trình diện vùng kháng nguyên H5 cũng được đánh giá. Catxet biểu hiện để biểu hiện sự dung hợp chất ngưng kết

hồng cầu H1/H5 được thể hiện trên Fig.1 và trình tự axit amin của protein dung hợp hoàn chỉnh tạo ra được thể hiện trên Fig.2.

Để so sánh mức tích tụ của chất ngưng kết hồng cầu thể khảm H1/H5 với mức tích tụ của dạng gốc của nó, các cây *Nicotiana benthamiana* được thâm nhiễm bởi AGL1/660 và AGL1/696, và các lá được thu sau khoảng thời gian sáu ngày ủ. Để xác định mức tích tụ của mỗi dạng HA trong các lá được thâm nhiễm vi khuẩn Ago, protein được chiết từ mô lá được thâm nhiễm và phân tích bởi thẩm tách Western sử dụng các kháng thể đa dòng khác H5 (Indonesia). Băng duy nhất khoảng 75 kDa (Fig.7), tương ứng với kích thước của dạng HA0 được phân tách của chất ngưng kết hồng cầu cúm, được phát hiện trong dịch chiết từ lá được thâm nhiễm bởi AGL1/660 và AGL1/696, chỉ ra rằng cả hai gốc H5 A/Indonesia/5/05 và chất ngưng kết hồng cầu thể khảm H1/H5 tích tụ ở mức cao trong cây.

Ví dụ 4: Sự thay thế vùng ngoài của H5 A/Indonesia/5/05 với vùng ngoài của H3 hoặc B. So sánh sự biểu hiện của các dạng khảm và dạng gốc.

Sự dung hợp ectodomain từ H3 A/Brisbane/10/2007 hoặc B Florida/4/2006 với các vùng phụ màng vận chuyển và tế bào chất từ H5 A/Indonesia/5/05 được đánh giá dưới dạng chiến dịch thể hiện các vùng khác nguyên chất ngưng kết hồng cầu từ các chủng H3 và B trong khi tăng mức tích lũy của chúng đối với các cây. Catxet biểu hiện để biểu hiện các điểm dung hợp chất ngưng kết hồng cầu H5/H3 và H5/B được thể hiện trên Fig.10 và axit amin tại đường viền của các điểm dung hợp được thể hiện trên Fig.11.

Mức tích tụ của chất ngưng kết hồng cầu thể khảm H5/B (745) được so sánh với gốc HA B (739) trong các cây *Nicotiana benthamiana*. Các cây *N. benthamiana* được thâm nhiễm AGL1/745, và thu được các lá sau giai đoạn 6 ngày ủ bệnh. Để xác định liệu H5 bị tích tụ trong các lá đã bị thâm nhiễm agro, protein được chiết lần thứ nhất từ mô lá đã được thâm nhiễm và được phân tích bằng phương pháp Western sử dụng các kháng thể đa dòng kháng-H5 (Vietnam). Băng duy nhất khoảng 75kDa (Fig.14), tương ứng với kích thước của dạng HA0 không được phân cắt của chất ngưng kết hồng cầu cúm, được phát hiện trong dịch chiết của từ lá của một cây được thâm nhiễm bởi AGL1/739 trong khi các cây 3 được thâm nhiễm bởi AGL1/745 chỉ ra tín hiệu dương tương ứng với chất ngưng kết hồng cầu, chỉ ra rằng dạng khảm H5/B của chất ngưng kết hồng cầu được tích tụ đều hơn tại mức cao hơn dạng gốc của chất ngưng kết hồng cầu B.

Tương tự, mức tích tụ chất ngưng kết hồng cầu thể khảm H5/H3 (737) được so sánh với dạng gốc của nó (736) trong các cây *Nicotiana benthamiana*. Các cây được thâm nhiễm bởi AGL1/736 và AGL1/737, và các lá được thu sau khoảng thời gian ủ sáu ngày. Để xác định liệu H5 bị tích tụ trong các lá đã bị thâm nhiễm agro, protein được chiết lần thứ nhất từ mô lá đã được thâm nhiễm và được phân tích bằng phương pháp Western sử dụng các kháng thể đa dòng kháng-H5 (Vietnam). Băng duy nhất khoảng 75kDa (Fig.15), tương ứng với kích thước của dạng HA0 không được phân cắt của chất ngưng kết hồng cầu cúm, được phát hiện trong các dịch chiết từ lá được thâm nhiễm bởi AGL1/736 và AGL1/737. Kết quả này chỉ ra rằng sự dung hợp các vùng phụ của màng vận chuyển và tế bào chất từ H5 A/Indonesia/5/05 đến vùng ngoài của H3 A/Brisbane/10/2007 tạo ra chất ngưng kết hồng cầu thể khảm tích tụ mức tương tự với H3 A/Brisbane/10/2007 gốc.

Sự sản xuất các VLP từ sự biểu hiện chất ngưng kết hồng cầu thể khảm H5/B (sản phẩm thiết kế số 745) được đánh giá sử dụng phương pháp sắc ký rây phân tử. Các dịch chiết protein thô từ các cây được thâm nhiễm AGL1/745 (1,5mL) được phân đoạn bằng sắc ký loại trừ kích thước (SEC) trên các cột Sephacryl™ S-500 HR (GE Healthcare Bio-Science Corp., Piscataway, NJ, USA). Như được chỉ ra trên Fig.16, sự rửa giải Blue Dextran (2MDa) được tạo đỉnh sớm trong đoạn 8. Khi các protein từ 200µL của mỗi phần rửa giải SEC được cô đặc (gấp 5 lần) bởi sự kết tủa acetone và được phân tích bằng thẩm tách Western sử dụng các kháng thể đa dòng kháng B (Florida) (Fig.16), chất ngưng kết hồng cầu thể khảm chủ yếu được tìm thấy trong đoạn 7, chỉ ra sự kết hợp của HA vào cấu trúc có trọng lượng phân tử lớn. Không mong muốn bị ràng buộc bởi lý thuyết, điều này đề xuất rằng protein HA được lắp vào trong sản phẩm thiết kế bên trên lớn hoặc là nó đã được gắn với sản phẩm thiết kế có trọng lượng phân tử cao. Các kết quả thu được chỉ ra rằng HA thể khảm gồm trong vùng ngoài từ HA B/Florida/4/2006 được dung hợp với các vùng phụ màng vận chuyển và xytosol từ H5 A/Indonesia/5/05 ghép vào các hạt có trọng lượng phân tử cao và đặc điểm rửa giải của các hạt có trọng lượng phân tử cao không thể phân biệt được với hạt của các VLP cúm.

Ví dụ 5: Sự biểu hiện đồng thời chất ngưng kết hồng cầu thể khảm H5/B (sản phẩm thiết kế số 747; bao gồm B/Flo HDC và SDC được dung hợp với H5/Indo TDC) với Hsp70 và Hsp40 kết hợp với sự biến đổi peptit tín hiệu

Sự biểu hiện Hsp40 và Hsp70 trong cây và biểu hiện đồng thời với chất ngưng kết hồng cầu cúm được mô tả trong đơn PCT/CA2009/000032 đang xử lý. Hsp70 và Hsp40 xytosol (sản phẩm thiết kế số R870) của cây có nguồn gốc có thể được biểu hiện đồng thời với chất ngưng kết hồng cầu thể khảm, làm tăng mức tích tụ của chúng trong cây. Sự biểu hiện đồng thời có thể được tiến hành bởi sự thâm nhiễm vi khuẩn *Agro* của các cây *N. benthamiana* với sự tạo huyền phù vi khuẩn chứa hỗn hợp (tỷ lệ 1:1:1) của AGL1 mang catxet để biểu hiện HA thể khảm mong muốn với AGL1/R870 và AGL1/35SHcPro.

Mức tích tụ của chất ngưng kết hồng cầu thể khảm H5/B (B/Flo HDC và SDC được dung hợp bởi H5/Indo TDC) được đánh giá trong sự biểu hiện đồng thời với HSP40 và HSP70 trong cây *Nicotiana benthamiana*. Các cây *N. benthamiana* được thâm nhiễm AGL1/747, AGL1/747, AGL1/443, AGL1/776, AGL1/747, AGL1/778 và AGL1/779 và các lá được thu hoạch sau giai đoạn 6 ngày ủ. Để xác định liệu H5 bị tích tụ trong các lá đã bị thâm nhiễm *agro*, protein được chiết lần thứ nhất từ mô lá đã được thâm nhiễm và được phân tích bằng phương pháp Western sử dụng các kháng thể đa dòng kháng-H5 (Vietnam). Băng duy nhất khoảng 75kDa (Fig.50), tương ứng với kích thước của dạng HA0 không được phân cắt của chất ngưng kết hồng cầu cúm, được phát hiện trong các dịch chiết từ lá từ 3 cây được thâm nhiễm bởi AGL1/747+AGL1/R870 trong khi 3 cây được thâm nhiễm bởi AGL1/747+ vector đối chứng (443) được chỉ ra không có tín hiệu (dưới điều kiện phơi nhiễm được sử dụng) chỉ ra rằng dạng thể khảm H5/B của chất ngưng kết hồng cầu được tích tụ với mức cao khi được biểu hiện đồng thời bởi các chaperon HSP40 và HSP 70.

Tất cả các chỉ dẫn được kết hợp ở đây bằng cách tham khảo, như thể mỗi đơn riêng lẻ được chỉ ra cụ thể và riêng lẻ được kết hợp bằng cách tham khảo ở đây và dường như đầy đủ được đưa ra trong bản mô tả. Sự trích dẫn các tài liệu tham khảo ở đây không được phân tích hoặc không được xem xét như sự thừa nhận rằng các tài liệu tham khảo này nằm trong tình trạng kỹ thuật của sáng chế.

Trong bản mô tả số lượng các thuật ngữ được sử dụng rộng rãi và các định nghĩa được đề xuất để hiểu dễ dàng hơn các khía cạnh khác nhau của sáng chế. Sử dụng các ví dụ trong bản mô tả, gồm các ví dụ về thuật ngữ, chỉ với mục đích minh họa và không được dùng để giới hạn phạm vi và nghĩa của các phương án thực hiện của sáng chế. Các khoảng giá trị bao gồm các giá trị định rõ khoảng. Trong bản mô tả, từ “bao gồm” được

sử dụng dưới dạng thuật ngữ mở rộng được, cơ bản tương đương với cụm từ “gồm, nhưng không giới hạn tới,” và từ “bao gồm” có nghĩa tương đương.

Một hoặc nhiều phương án thực hiện đang được ưu tiên theo sáng chế được mô tả bằng ví dụ. Sáng chế gồm tất cả các phương án thực hiện, các cải biến và biến thể cơ bản được mô tả ở trên và về các ví dụ và Fig.. Tuy nhiên, sẽ rõ ràng đối với người có trình độ trung bình trong lĩnh vực kỹ thuật này rằng những biến đổi và cải biến có thể được tạo ra mà không tách rời khỏi phạm vi của sáng chế như được xác định trong các yêu cầu bảo hộ. Các ví dụ về các cải biến này gồm sự thay thế các cải biến tương đương đã biết trong khía cạnh bất kỳ của sáng chế để đạt được kết quả tương tự cơ bản theo cách tương tự.

YÊU CẦU BẢO HỘ

1. Axit nucleic chứa một hoặc nhiều vùng điều hòa nhận

II hoạt động ở thực vật và được liên kết hoạt động với trình

thể khảm bao gồm nhóm vùng gốc (SDC), nhóm vùng đầu (HDC) và nhóm vùng xuyên màng (TDC), một hoặc nhiều vùng điều hòa này còn bao gồm thêm 5'UTR, 3'UTR hoặc 5'UTR và 3'UTR mà hoạt động ở thực vật, côn trùng hoặc nấm men, và trong đó:

a) SDC bao gồm vùng phụ F'1, F'2 và F;

b) HDC bao gồm vùng phụ RB, E1 và E2;

c) TDC bao gồm vùng phụ TmD và Ctail; và

trong đó SDC và HDC là thuộc HA cúm thứ nhất và TDC thuộc một hoặc nhiều HA cúm thứ hai, trong đó HA cúm thứ nhất và một hoặc nhiều HA cúm thứ hai thu được từ các chủng cúm khác nhau.

2. Axit nucleic theo điểm 1, trong đó trình tự mã hóa polypeptit HA cúm thể khảm còn bao gồm trình tự peptit tín hiệu được chọn từ nhóm gồm trình tự peptit tín hiệu tự nhiên HA, trình tự peptit tín hiệu alfalfa PDI, trình tự peptit tín hiệu H5 cúm và trình tự peptit tín hiệu H1 cúm.

3. Axit nucleic theo điểm 1, trong đó 5'UTR, 3'UTR hoặc 5'UTR và 3'UTR thu được từ plastoxyanin UTR hoặc CPMV UTR.

4. Axit nucleic theo điểm 1, trong đó vùng điều hòa thu được từ vùng điều hòa plastoxyanin, vùng điều hòa ribuloza 1,5-bisphosphat carboxylaza/oxygenaza (RuBisCO), vùng điều hòa protein gắn chlorophyll a/b (CAB), vùng điều hòa CaMV 35S, vùng điều hòa actin, vùng điều hòa ubiquitin, vùng điều hòa triozaphosphat isomeraza 1, vùng điều hòa yếu tố khởi đầu dịch mã 4A, và vùng điều hòa ST-LSI.

5. Axit nucleic theo điểm 1, trong đó HA cúm thứ nhất và thứ hai độc lập được chọn từ nhóm gồm H1, H3, H5 và B.

6. Phương pháp sản xuất hạt tương tự virus cúm thể khảm (VLP) ở thực vật bao gồm bước:

a) đưa axit nucleic theo điểm 1 vào cây, hoặc bộ phận của cây, và

b) ủ cây, hoặc bộ phận của cây trong các điều kiện cho phép biểu hiện axit nucleic này, bằng cách đó tạo ra các VLP.

7. Phương pháp theo điểm 6, trong đó ở bước đưa vào (bước a), axit nucleic được đưa vào cây theo kiểu tạm thời.

8. Phương pháp theo điểm 6, trong đó, ở bước đưa vào (bước a), axit nucleic được đưa vào cây sao cho nó được kết hợp một cách ổn định vào hệ gen.

9. Phương pháp theo điểm 6, trong đó phương pháp này còn bao gồm bước:

c) thu hoạch vật chủ và tinh chế các VLP.

10. Polypeptit được mã hóa bởi axit nucleic theo điểm 1.

11. Hạt tương tự virus (VLP) chứa polypeptit theo điểm 10.

12. Hạt tương tự virus theo điểm 11, trong đó hạt này còn chứa thêm các N-glycan đặc hiệu thực vật hoặc các N-glycan được cải biến.

13. Chế phẩm vacxin chứa lượng hữu hiệu hạt tương tự virus (VLP) theo điểm 12 và chất mang được dụng.

14. Cây chứa polypeptit được mã hóa bởi axit nucleic theo điểm 1.

15. Bộ phận của cây chứa polypeptit được mã hóa bởi axit nucleic theo điểm 1.

16. Axit nucleic chứa một hoặc nhiều vùng điều hòa được liên kết hoạt động với trình tự mã hóa polypeptit HA cúm thể khảm chứa nhóm vùng gốc (SDC), nhóm vùng đầu (HDC) và nhóm vùng xuyên màng (TDC), trong đó một hoặc nhiều vùng điều hòa này bao gồm 5'UTR, 3'UTR hoặc 5'UTR và 3'UTR, và trong đó

a) SDC bao gồm vùng phụ F'1, F'2 và F;

b) HDC bao gồm vùng phụ RB, E1 và E2;

c) TDC bao gồm vùng phụ TmD và Ctail; và

trong đó vùng phụ RB thuộc HA cúm thứ nhất và vùng phụ E1 và E2 thuộc một hoặc nhiều HA cúm thứ hai, và trong đó SDC hoặc TDC hoặc cả SDC và TDC thuộc một hoặc nhiều HA cúm thứ hai, trong đó HA cúm thứ nhất và một hoặc nhiều HA cúm thứ hai thu được từ các chủng cúm khác nhau.

17. Axit nucleic theo điểm 16, trong đó E1, E2 hoặc cả hai vùng phụ này và F'1, F'2 hoặc cả hai vùng phụ này thuộc cùng một HA cúm.
18. Polypeptit được mã hóa bởi axit nucleic theo điểm 16.
19. Hạt tương tự virus (VLP) chứa polypeptit theo điểm 18.
20. Phương pháp tạo ra hạt tương tự virus cúm thể khảm (VLP) ở thực vật bao gồm bước:
- a) đưa axit nucleic theo điểm 18 vào cây, hoặc bộ phận của cây, và
 - b) ủ cây, hoặc bộ phận của cây trong các điều kiện cho phép biểu hiện axit nucleic này, bằng cách đó tạo ra các VLP.

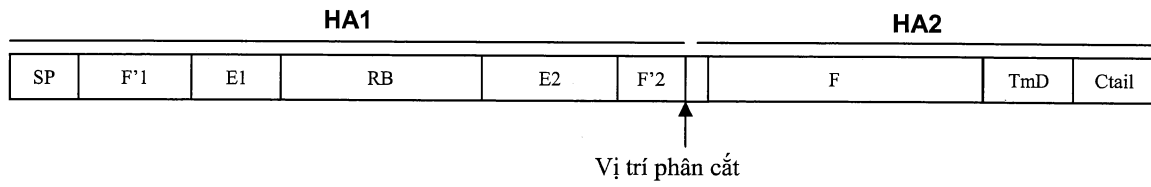


Fig. 1a

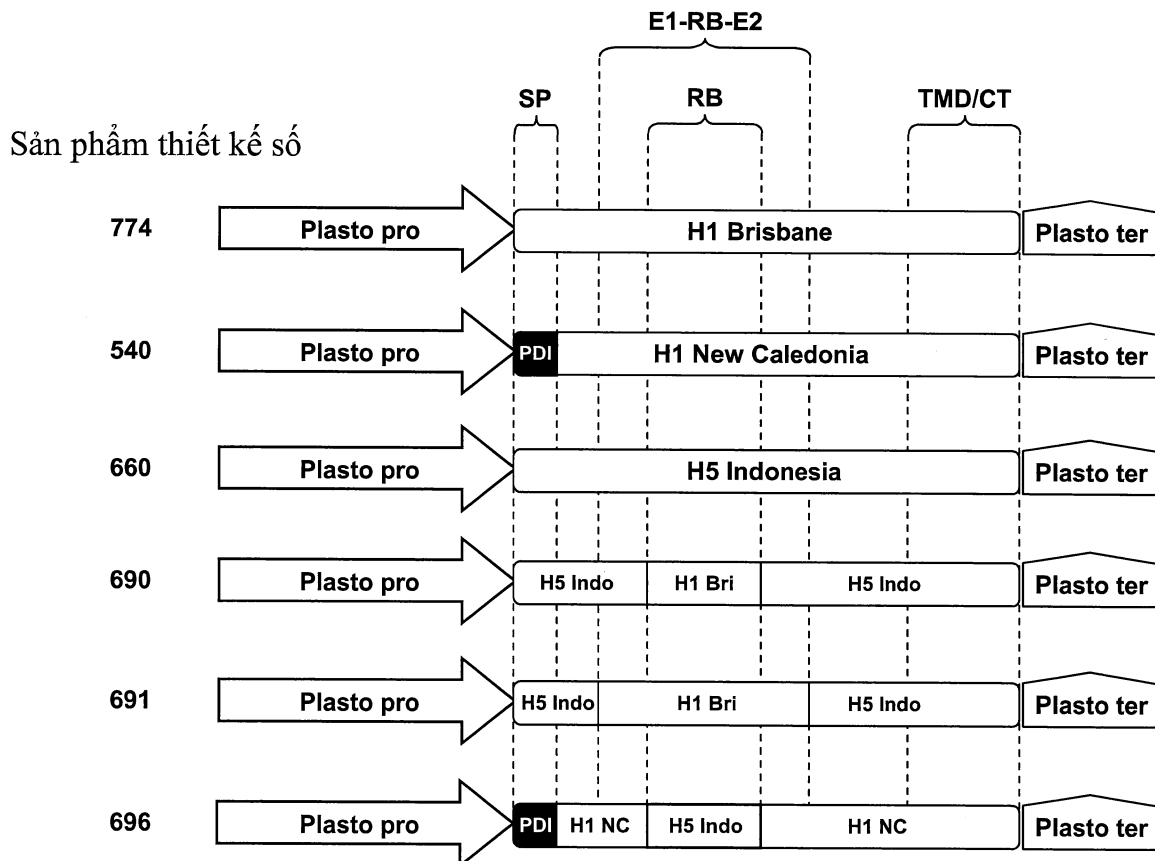


Fig. 1b

2/49

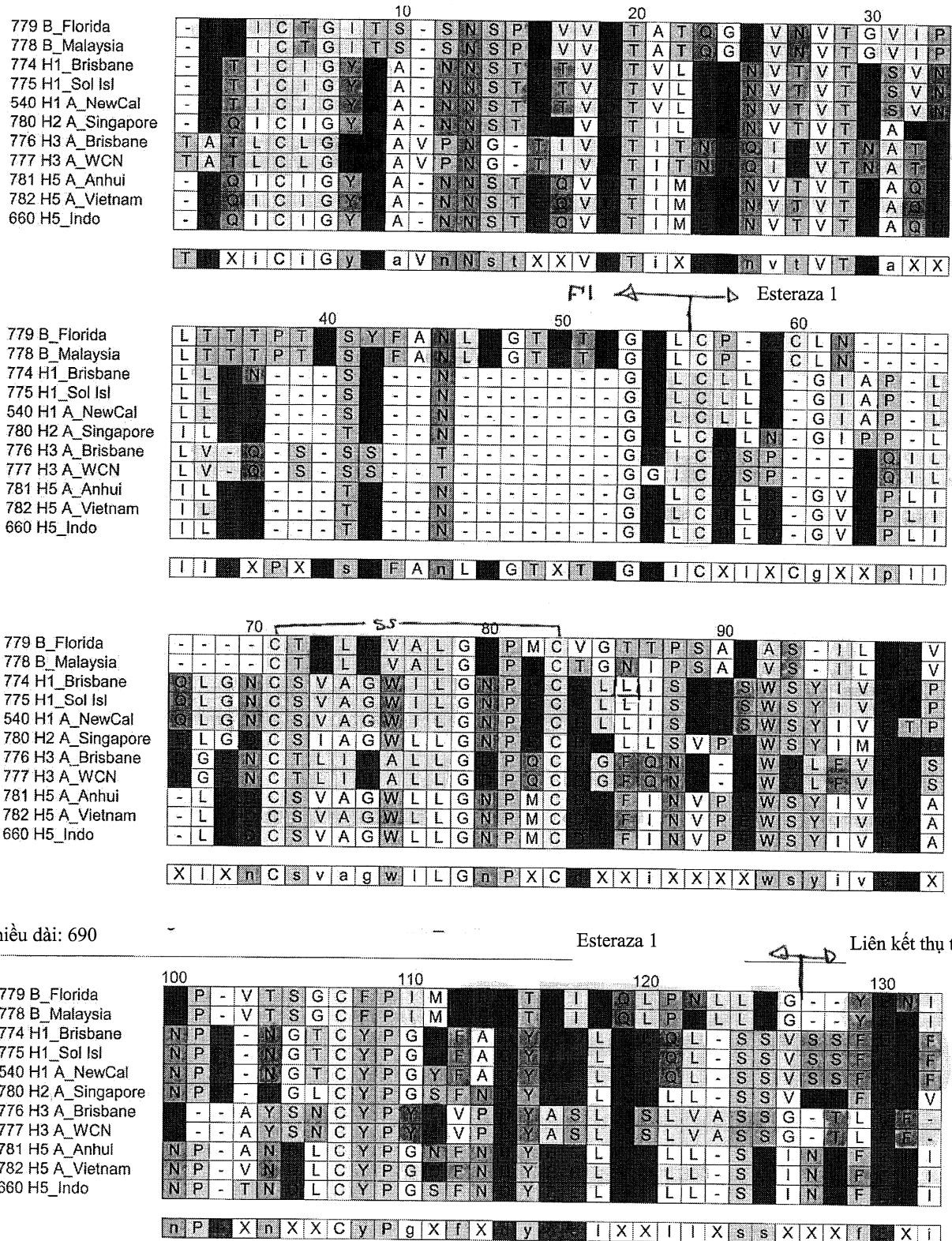


Fig. 1C-1

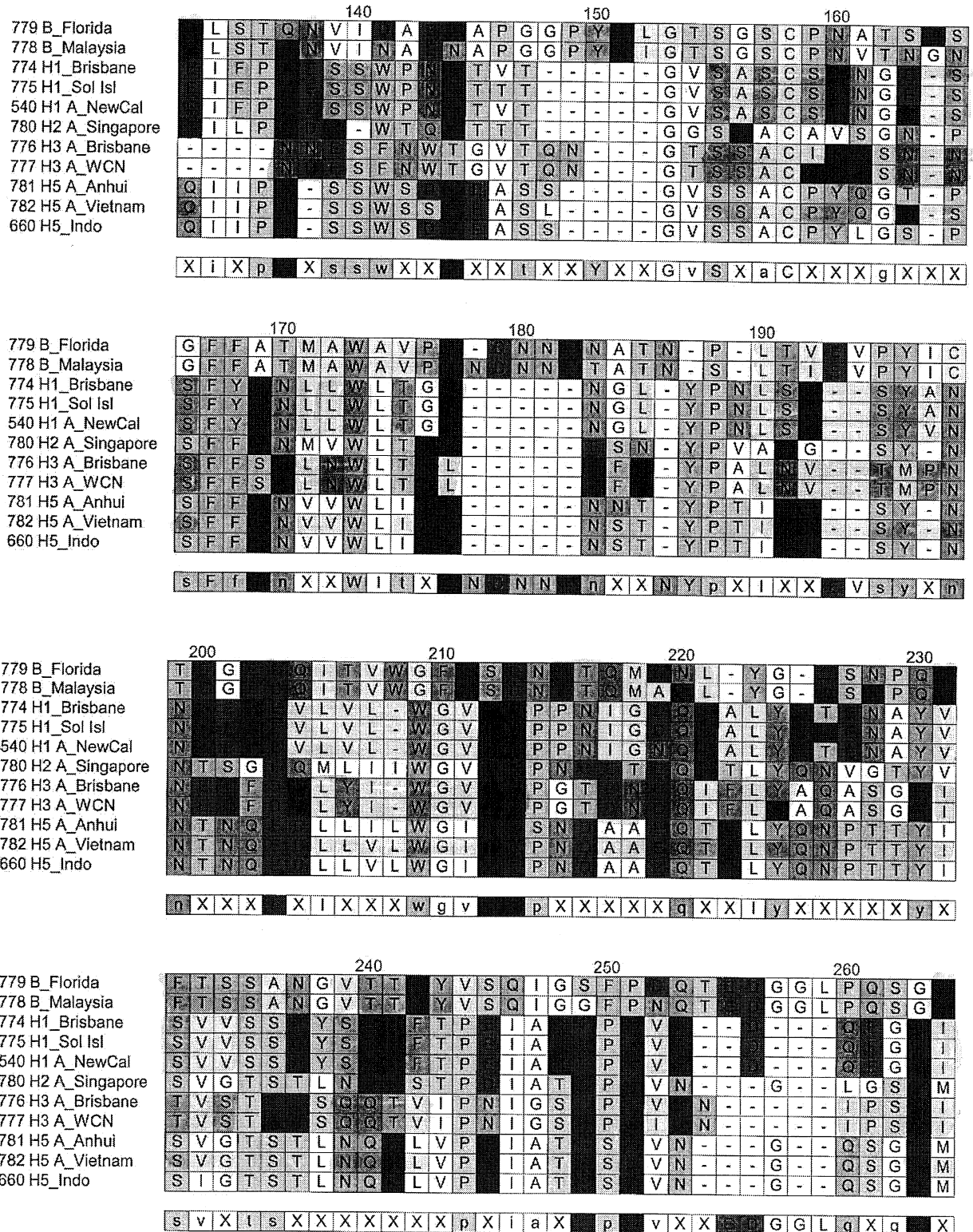


Fig. 1C-2

4/49

779 B_Florida
778 B_Malaysia
774 H1_Brisbane
775 H1_Sol Isl
540 H1_A_NewCal
780 H2_A_Singapore
776 H3_A_Brisbane
777 H3_A_WCN
781 H5_A_Anhui
782 H5_A_Vietnam
660 H5_Indo

270	280	290
I V V Y M M Q P G T G T I V Y Q G V L L P Q V W C A S -	I V V Y M V Q S G T G T I T Y Q G I L L P Q V W C A S -	
N Y Y W T L L P G T I I F A N G N - L I A P Y A F A L S	N Y Y W T L L P G T I I F A N G N - L I A P Y A F A L S	
N Y Y W T L L P G T I I F A N G N - L I A P W Y A F A L S	N Y Y W T L L P G T I I F A N G N - L I A P W Y A F A L S	
F S W T L L M W T I N F S T G N - L I A P Y G F I S	F S W T L L M W T I N F S T G N - L I A P Y G F I S	
S I Y W T I V P G I L L I N S T G N - L I A P G Y F I -	S I Y W T I V P G I L L I N S T G N - L I A P G Y F I -	
F F W T I L P N A I N F S N G N - F I A P Y A Y I V	F F W T I L P N A I N F S N G N - F I A P Y A Y I V	
F F W T I L P N A I N F S N G N - F I A P Y A Y I V	F F W T I L P N A I N F S N G N - F I A P Y A Y I V	
X X X w t X I X p X X i X f s n g n G l i a P X y a f i s		

Chiều dài: 609
Liên kết thụ thể

← Esteraza 2 → F'2

779 B_Florida
778 B_Malaysia
774 H1_Brisbane
775 H1_Sol Isl
540 H1_A_NewCal
780 H2_A_Singapore
776 H3_A_Brisbane
777 H3_A_WCN
781 H5_A_Anhui
782 H5_A_Vietnam
660 H5_Indo

300	310	320	330
- G S V I G S L P L I G - A C L Y G G L N S P	- G S V I G S L P L I G - A C L Y G G L N S P		
G F G S G I I N S N A P M C A C G T P Q - G A I N S S L P	G F G S G I I N S N A P M C A C G T P Q - G A I N S S L P		
G F G S G I I N S N A P M C A C G T P Q - G A I N S S L P	G F G S G I I N S N A P M C A C G T P Q - G A I N S S L P		
G S S G I M T G T L N C T C Q T P L - G A I N T T L P	G S S G I M T G T L N C T C Q T P L - G A I N T T L P		
S G S S I M S A P I G C N S C I T P N - G S I P N P	S G S S I M S A P I G C N S C I T P N - G S I P N P		
S G S S I M S A P I G C N S C I T P N - G S I P N P	S G S S I M S A P I G C N S C I T P N - G S I P N P		
G S A I V S V Y G N C N T C Q T P I - G A I N S S M P	G S A I V S V Y G N C N T C Q T P I - G A I N S S M P		
G S A I M S L Y G N C N T C Q T P M - G A I N S S M P	G S A I M S L Y G N C N T C Q T P M - G A I N S S M P		
X g X S X i X s X X p X X X C X X C g t p X Y G a i n s s X P			

779 B_Florida
778 B_Malaysia
774 H1_Brisbane
775 H1_Sol Isl
540 H1_A_NewCal
780 H2_A_Singapore
776 H3_A_Brisbane
777 H3_A_WCN
781 H5_A_Anhui
782 H5_A_Vietnam
660 H5_Indo

340	350	360
Y Y T G A A I G N C P I W V T P - L L A N G T Y P -	Y Y T G A A I G N C P I W V T P - L L A N G T Y P -	
F Q N V - P V T I G C P Y V S A L M V - - T G L N I	F Q N V - P V T I G C P Y V S A L M V - - T G L N I	
F Q N V - P V T I G C P Y V S A L M V - - T G L N I	F Q N V - P V T I G C P Y V S A L M V - - T G L N I	
F Q N V - P V T I G C P Y V S A L M V - - T G L N I	F Q N V - P V T I G C P Y V S A L M V - - T G L N I	
F Q N V - N I T Y G A C P Y V Q N T L L A - - T G M N V	F Q N V - N I T Y G A C P Y V Q N T L L A - - T G M N V	
F Q N V - N I T Y G A C P Y V Q N T L L A - - T G M N V	F Q N V - N I T Y G A C P Y V Q N T L L A - - T G M N V	
F N I - P L T I G C P Y V S N L V L A - - T G L N S	F N I - P L T I G C P Y V S N L V L A - - T G L N S	
F N I - P L T I G C P Y V S N L V L A - - T G L N S	F N I - P L T I G C P Y V S N L V L A - - T G L N S	
f X n v p X t i G C P y V s X L X i a N G T g l n X		

F'2 H1 ← H2 Dung hợp peptit

779 B_Florida
778 B_Malaysia
774 H1_Brisbane
775 H1_Sol Isl
540 H1_A_NewCal
780 H2_A_Singapore
776 H3_A_Brisbane
777 H3_A_WCN
781 H5_A_Anhui
782 H5_A_Vietnam
660 H5_Indo

370	380	390
P A L - - L G F F G A I A G F L G G W G M I A G W G	P A L - - L G F F G A I A G F L G G W G M I A G W G	
P S - - - I Q S G L F G A I A G F I G G W T G M V G W Y G	P S - - - I Q S G L F G A I A G F I G G W T G M V G W Y G	
P S - - - I Q S G L F G A I A G F I G G W T G M V G W Y G	P S - - - I Q S G L F G A I A G F I G G W T G M V G W Y G	
P Q - - - I S G L F G A I A G F I G G W Q G M V G W Y G	P Q - - - I S G L F G A I A G F I G G W Q G M V G W Y G	
P - - - - Q T G I F G A I A G F I N G W G M V G W Y G	P - - - - Q T G I F G A I A G F I N G W G M V G W Y G	
P L - - - - G L F G A I A G F I G G W Q G M V G W Y G	P L - - - - G L F G A I A G F I G G W Q G M V G W Y G	
P Q - - - - G L F G A I A G F I G G W Q G M V G W Y G	P Q - - - - G L F G A I A G F I G G W Q G M V G W Y G	
P Q - - - - G L F G A I A G F I G G W Q G M V G W Y G	P Q - - - - G L F G A I A G F I G G W Q G M V G W Y G	
P X X X X X X X G I F G A I A G F i g G W X G M v G W y G		

Fig. 1C-3

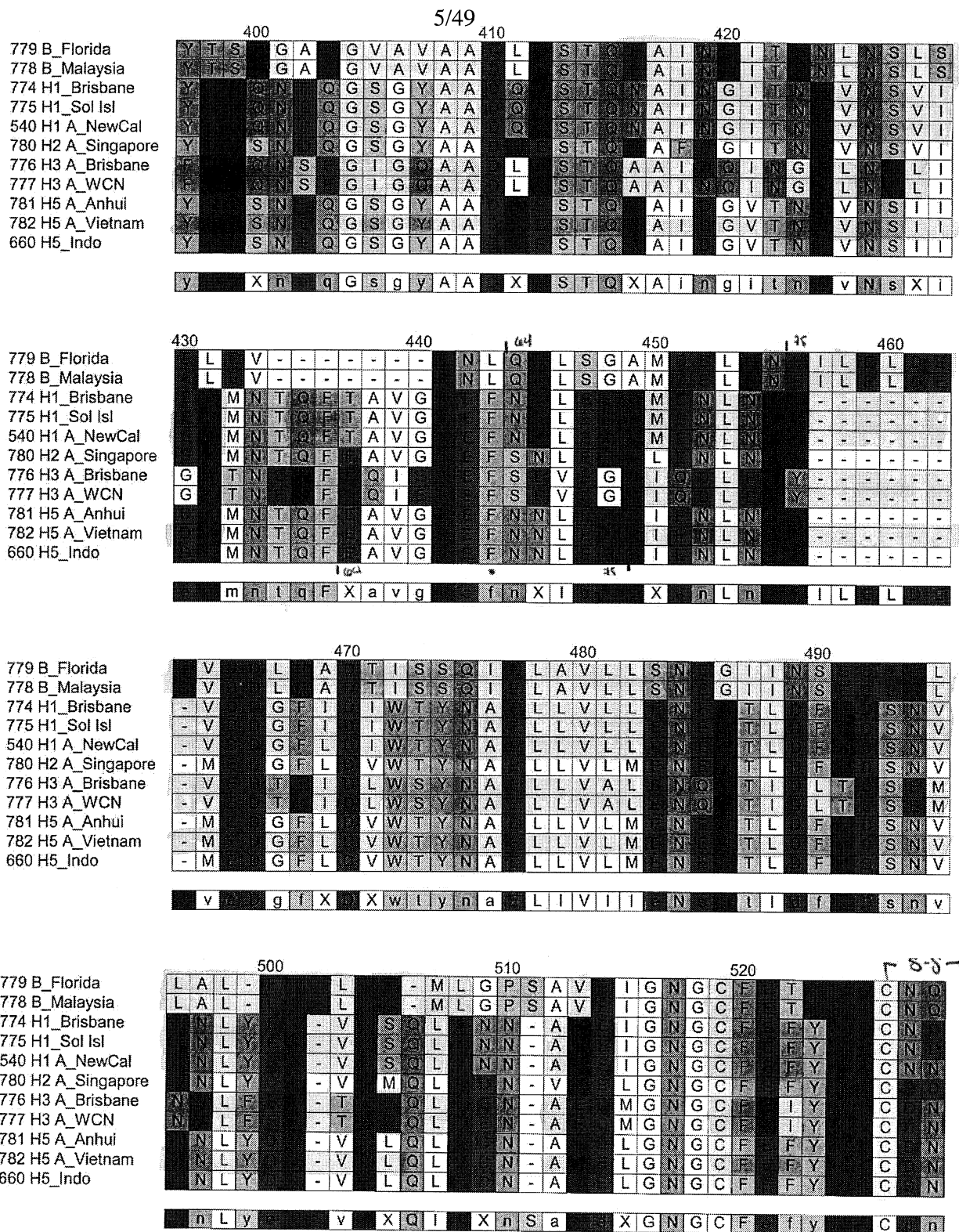


Fig. 1C-4

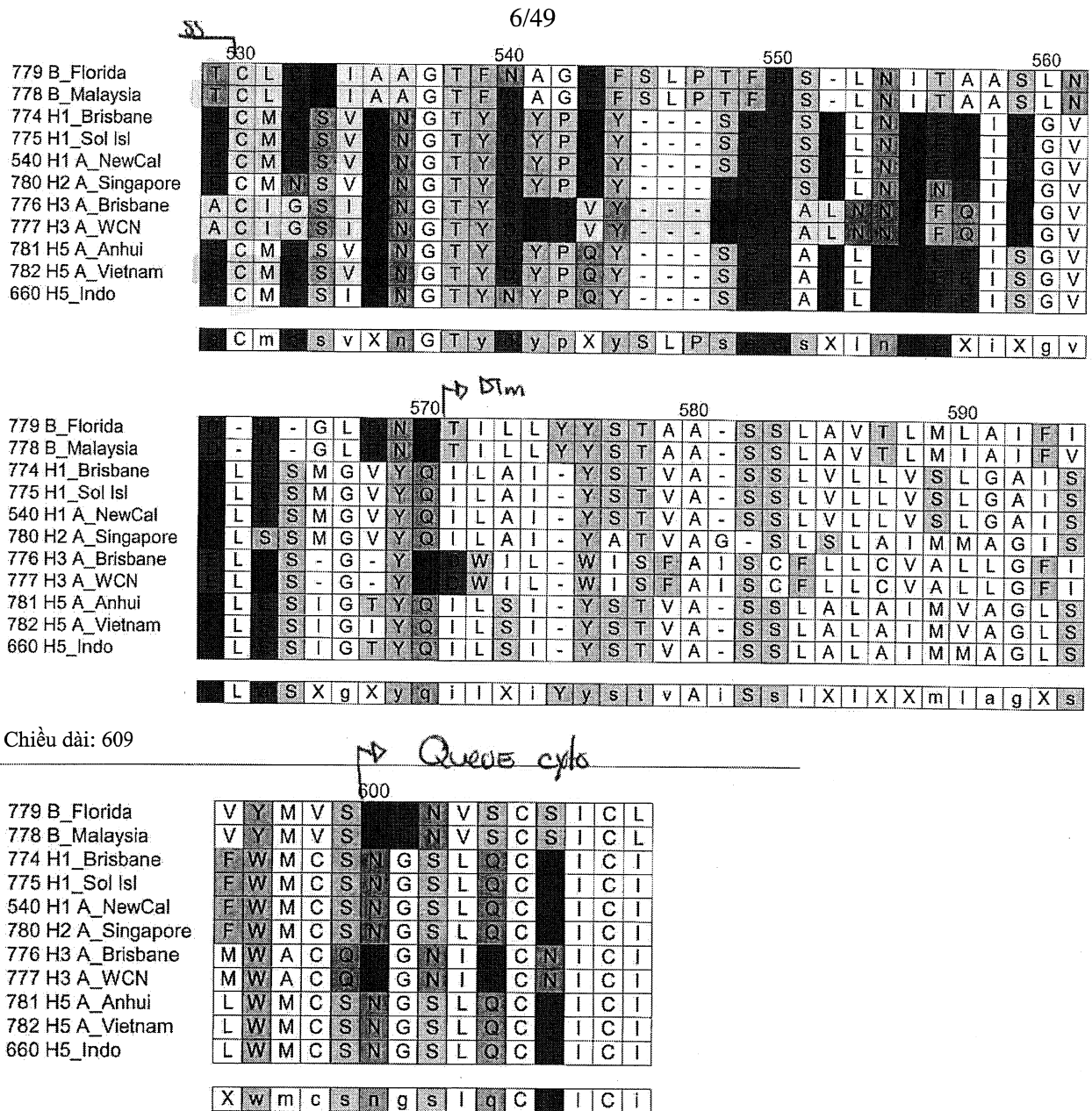


Fig. 1C-5

Fig. 2

Hoán đổi vùng phụ liên kết thụ thể trên HA gốc. Trình tự của protein hoàn chỉnh

Sản phẩm thiết kế số	Đầu tận cùng N gốc:- F'1+E1	Đầu RB	Đầu tận cùng C gốc: E2+F'2-Kết thúc
690 và 734	H5/Indo DQICIGYHANNS TEQVDTIMEKNV TVTHAQDILEKT HNGKLCDLG KPLILRDCSVAG WLLGNPMCDEF INVPEWSYIVEK ANPTNDLCYP	H1/Brisbane GHFADYEELREQLSSVSSFERFEIF PKESSWPNTVTGVSASCSHNGES SFYRNLLWLTGKNGLYPNLSKSYA NNKEKEVLVLWGVIHPPNIGDQKA LYHTENAYVSVSSHYSRKFTPEIA KRPKVRDQEGRINYWTLLPGDTI IFEANGNLIAPRYAFALSRGFGSG	H5/Indo IMKSELEYGNCNTKCQTPMGAINSSM PFHNIHPLTIGECPKYVKSRLVLATGL RNSPQRESRRKKRGLFGAIGFIEGG WQGMVDGWYGYHHSNEQGSYAAD KESTQKAIDGVTNKVNSIIDKMNTQFE AVGREFNLERRIENLNKKMEDGFLDV WTYNAELLVLMENERTLDFHDSNVKN LYDKVRLQLRDNAKELGNGCFEFYHK CDNECMESIRNGTYNYPQYSEEARLK REEISGVKLESIGTYQILSIYSTVASSLA LAIMMAGLSLWMCNGLQCRICI
696	H1/NC DTICIGYHANNS TDTVDTVLEKNV TVTHSVNLLDS HNGKLCCLKGIA PLQLGNCSVAG WILGNPECELLI SKESWSYIVETP NPENGTCTYP	H5/Indo GSFNDYEELKHLLSRINHFEKIQIIP KSSWSDHEASSGVSSACPYLGSPS FFRNVVWLIKKNSTYPTIKKSYNNT NQEDLLVLWGIHHPNDAAEQTRLY QNPTTYISIGTSTLNQRLVPKIATRS KVNGQSGRMEFFWTILKPNDAINF SNGNFIAPRYAYKIVKKGDASA	H1/NC IITSNAPMDECDKQTPQGAINSSLP FQNVHPVTIGECPKYVRSALRMVTGL RNIPSIQSRGLFGAIGFIEGGWGTGMV DGWYGYHHQNEQGSYAADQKSTQ NAINGITNKVNSVIEKMNTQFTAVGKEF NKLERRMENLNKKVDDGFLDIWTYNA ELLVLENERTLDFHDSNVKNLYEKVK SQLKNNAKEIGNGCFEFYHKCNNECM ESVKNGTYPKYSEESKLNREKIDGV KLESMGVYQILAIYSTVASSLVLLVSLG AISFWMCNGLQCRICI

Hoán đổi vùng phụ liên kết thụ thể và esterase trên HA gốc.
Trình tự của protein hoàn chỉnh

Sản phẩm thiết kế số	Đầu tận cùng N gốc: F'1	Đầu E1-RB-E2	Đầu tận cùng C gốc: F'2-Kết thúc
691	H5/Indo DQICIGYHANNS TEQVDTIMEKNV TVTHAQDILEKT HNGKLC	H1/Brisbane LLKGIAPLQLGNCSVAGWIL GNPECELLISKESWSYIVEK PNPENGTCYPGHFADYEEL REQLSSVSSFERFEIFPKES SWPNHTVTGVSASCSHNG ESSFYRNLLWLTGKNGLYP NLSKSYANNKEKEVLVLWG VHPPNIGDQKALYHTENA YVSVSSHYSRKFTPEIAKR PKVRDQEGRINYWTLLP GDTIIFEANGNLIAPRYAFAL SRGFGSGIINSNAPMDK	H5/Indo CNTKCQTPMGAINSSMPFHNIHPLTIGECPKYVKS RLVLATGLRNSPQRESRRKKRGLFGAIGFIEGGW QGMVDGWYGYHHSNEQGSYAADKESTQKAIDG VTNKVNSIIDKMNTQFEAVGREFNLERRIENLNKK MEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNL YDKVRLQLRDNAKELGNGCFEFYHKCDNECMESIR NGTYNYPQYSEEARLKREEISGVKLESIGTYQILSIY STVASSLALAIMMAGLSLWMCNGLQCRICI

Fig. 3

Sự dịch mã axit amin của trình tự mã hóa trong cassette biểu hiện sản phẩm thiết kế 690 và 734 (SEQ ID NO: 80):

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTTHAQDILEKTHNGKLCDLGDGVKPLILR
DCSVAGWLLGNPMCMDEFINVPESYIVEKANPTNDLCYPGHFADYEELREQLSSVSSFERFEIFPK
ESSWPNHVTGVSASCSHNGESSFYRNLLWLTGKNGLYPNLSKSYANNKEKEVLVLWGVHHPNI
GDQKALYHTENAYVSVVSSHYSRKFTPEIAKRPKVRDQEGRINYWTLLPEGDTIIFEANGNLIAPRY
AFALSRGFGSGIMKSELEYGNCNTKCQTPMGAINSSMPFHNIHPLTIGECPKYVKS NRLVLATGLRN
SPQRESRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQSGSYAADKESTQKAIDGV TNKVN
SIIDKMNTQFEAVGREFNLERRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYD
KVRLQLRDNAKELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEISGVKLESIGTYQIL
SIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

Fig. 4

Sự dịch mã axit amin của trình tự mã hóa trong cassette biểu hiện sản phẩm thiết kế 691 (SEQ ID No: 81):

MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTTHAQDILEKTHNGKLCCLKGIAPLQL
GNCSVAGWILGNPECELLISKESWSYIVEKPNPENGTCYPGHFADYEELREQLSSVSSFERFEIF
PKESSWPNHVTGVSASCSHNGESSFYRNLLWLTGKNGLYPNLSKSYANNKEKEVLVLWGVH
PPNIGDQKALYHTENAYVSVVSSHYSRKFTPEIAKRPKVRDQEGRINYWTLLPEGDTIIFEANGN
LIAPRYAFALSRGFGSGIINSNAPMDKCNTKCQTPMGAINSSMPFHNIHPLTIGECPKYVKS NRLV
LATGLRN SPQRESRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQSGSYAADKESTQKA
IDGV TNKVN SIIDKMNTQFEAVGREFNLERRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDF
HDSNVKNLYDKVRLQLRDNAKELGNGCFEFYHKCDNECMESIRNGTYNYPQYSEEARLKREEIS
GVKLESIGTYQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

Fig. 5

Sự dịch mã axit amin của trình tự mã hóa trong cassette biểu hiện sản phẩm thiết kế 696 (SEQ ID NO: 82):

MAKNVAIFGLLFSLVLVPSQIFADTICIGYHANNSTDVDTVLEKNVTVTTHSVNLLDSHNGKLCCLKGIA
PLQLGNCSVAGWILGNPECELLISKESWSYIVETPNPENGTCYPGSFNDYEELKHLLSRINHFEKIQIIPK
SSWSDHEASSGVSSACPYLGSPSFFRN VVWLKKNSTYPTIKKSYNNNTNQEDLLVLWGIHHPNDAAEQT
RLYQNPTTYISIGTSTLNQRLVPKIATRSKVNQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKK
GDSAIITSNAPMDECDACQTPQGAINSSLPFQNVHPVTIGECPKYVRS AKLRMVTGLRNIPS IQSRGLF
GAIAGFIEGGWTGMVDGWYGYHHQNEQSGSYAADQKSTQNAINGITNKVNSVIEKMNTQFTAVGKEF
NKLERRMENLNKKVDDGFLDIWTYNAELLVLENER TLDFHDSNVKNLYEKVKS QLKNNAKEIGNGCFE
FYHKCNNECMESVKNGTYDYPKYSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFW
MCSNGSLQCRICI

9/49

Fig. 6

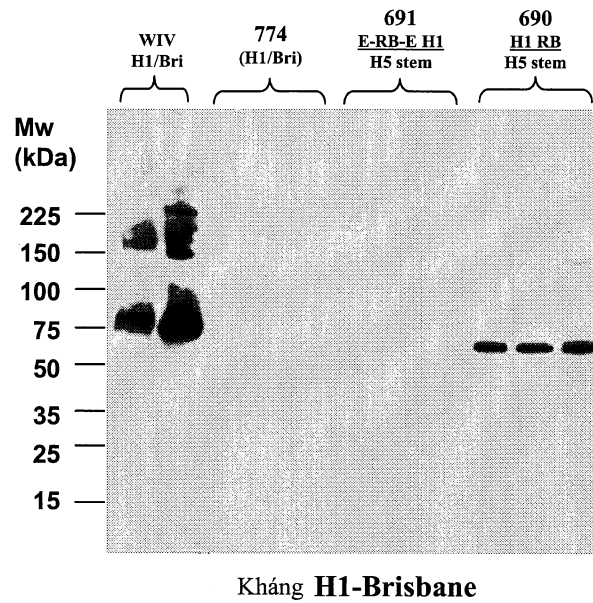


Fig. 7

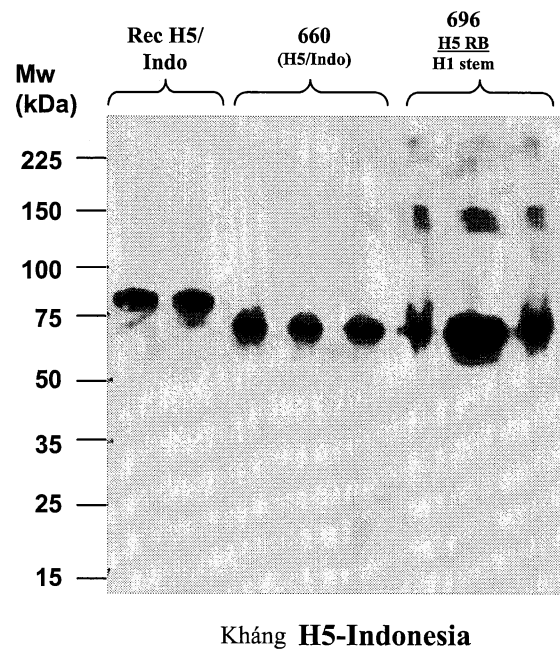


Fig. 8

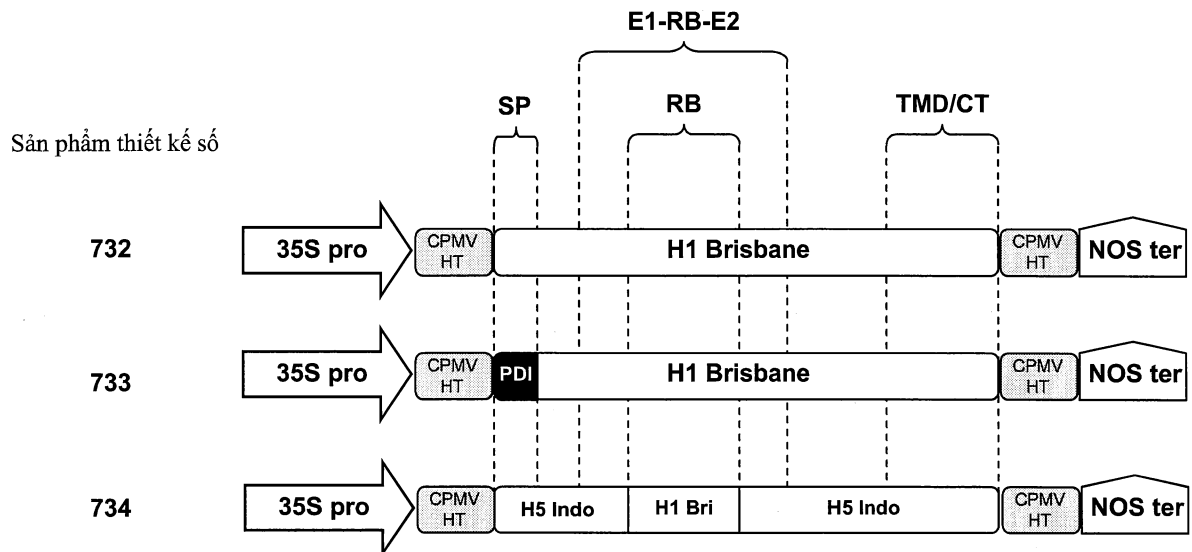
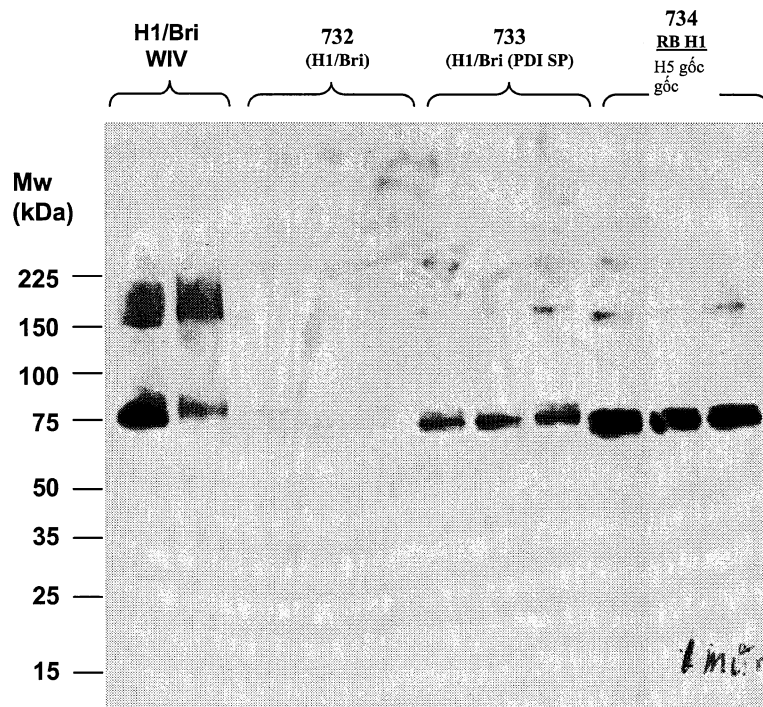


Fig. 9



Kháng H1-Brisbane

Fig. 10

Sản phẩm thiết kế số

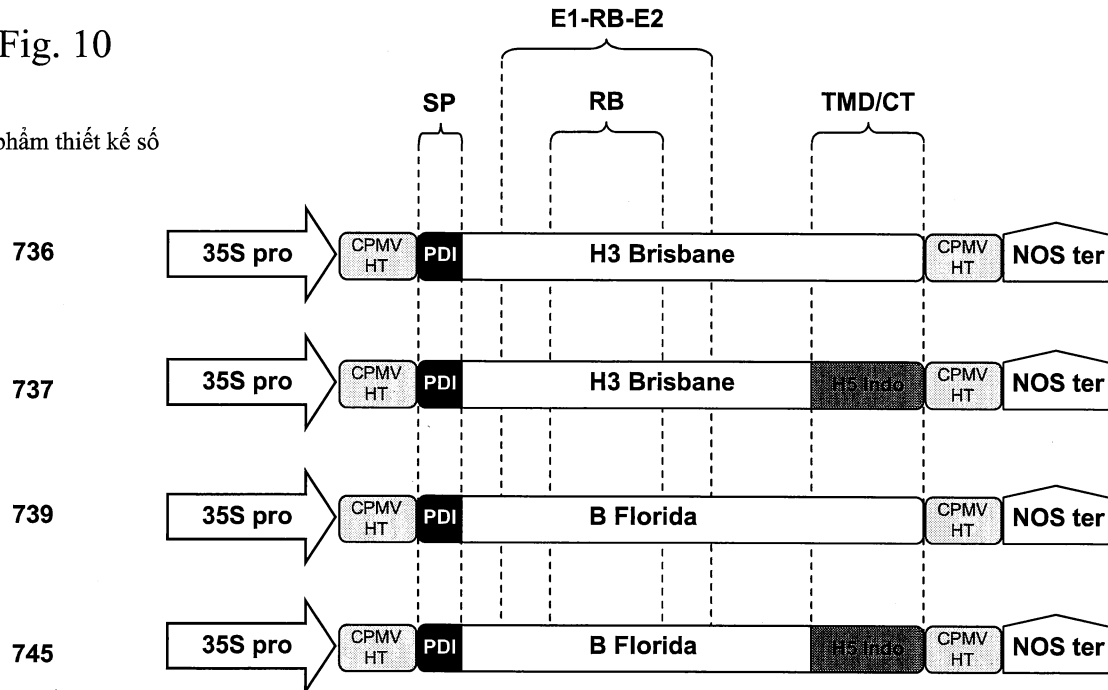


Fig. 11

Sản phẩm thiết kế số

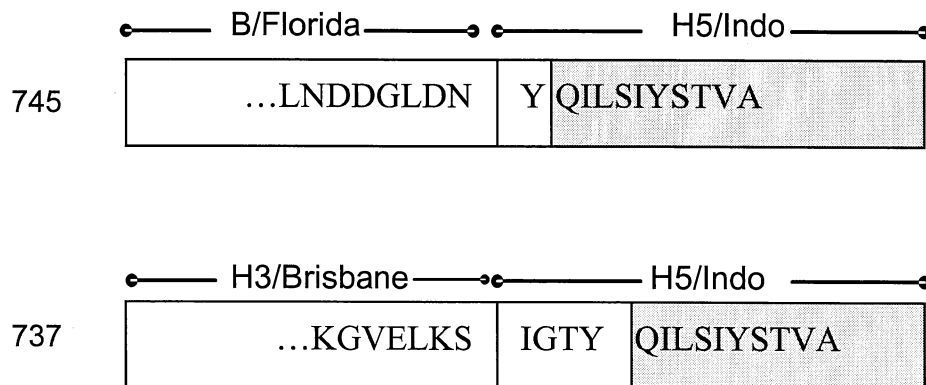


Fig. 12

Sự dịch mã axit amin của trình tự mã hóa trong cassette biểu hiện sản phẩm thiết kế 737 (SEQ ID NO: 83):

MAKNVAIFGLLFSLVLVPSQIFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSS
STGEICDSPHQILDGENCTLIDALLGDPQCDGFQNKWDLFVERSKAYSNCYPYDVPDYASLRSLVA
SSGTLEFNNEFNNWTGVTQNGTSSACIRRSNNSFFSRLNWLTHLKFKYPALNVTMPNNEKFDKLYIW
GVHHPGTDNDQIFLYAQASGRITVSTKRSQQTVIPNIGSRPRVRNIPSRISYWTIVKPGDILLINSTGNL
IAPRGYFKIRSGKSSIMRSDAIGKCNSECITPNGSIPNDKPFQNVNRITYGACPRYVKQNTLKLATGM
RNVPEKQTRGIFGAIAGFIENGWEGMVDGWYGRHQNSEGIGQAADLKSTQAAIDQINGKLNRLIGK
TNEKFHQIEKEFSEVEGRIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSMNKLFKTKKQLR
ENAEDMGNGCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSIGTYQILSIYSTVASS
LALAIMMAGLSLWMCSNGSLQCRICI

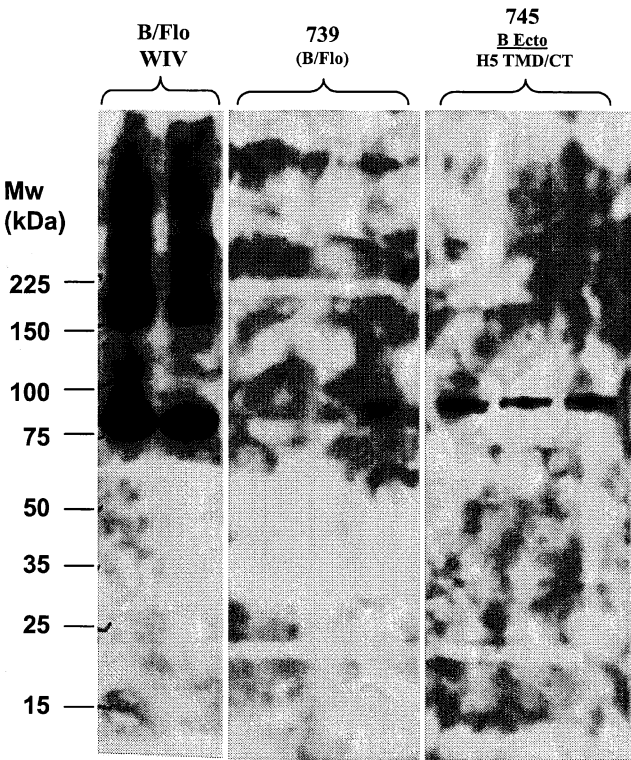
Fig. 13

Sự dịch mã axit amin của trình tự mã hóa trong cassette biểu hiện sản phẩm thiết kế 745 (SEQ ID NO: 84):

MAKNVAIFGLLFSLVLVPSQIFADRICTGITSSNSPHVVKATQGEVNVTVGVIPLTTTPTKSYFANLKGTR
TRGKLCPDCLNCTDLVALGRPMC VGTTPSAKASILHEVKPVTSGCFPIIMHDRTKIRQLPNLLRGYENI
RLSTQNVIDAEKAPGGPYRLGTSGSCPNATSKSGFFATMAWAVPKDNNKNATNPLTVEVPYICTEGED
QITVWGFHSDNKTQMKNLYGDSNPQKFTSSANGVTTHYVSQIGSFDPDQTEDGGLPQSGRIVVDYMMQ
KPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSPLIGEADCLHEKYGGLNKS KPYTGEHAKAIGNCP
IWVKTPLKLANGTKYRPPAKLLKERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQE
AINKITKNLNSLSELEVKNLQRLSGAMDELHNEILELDEKVDDL RADTISSQIELAVLLSNEGIINSEDEHLL
ALERKLLKMLGPSAVEIGNGCFETKHKCNQTC LDRIAAGTFNAGEFSLPTFD SLNITAASLNDDGLDNY
QILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

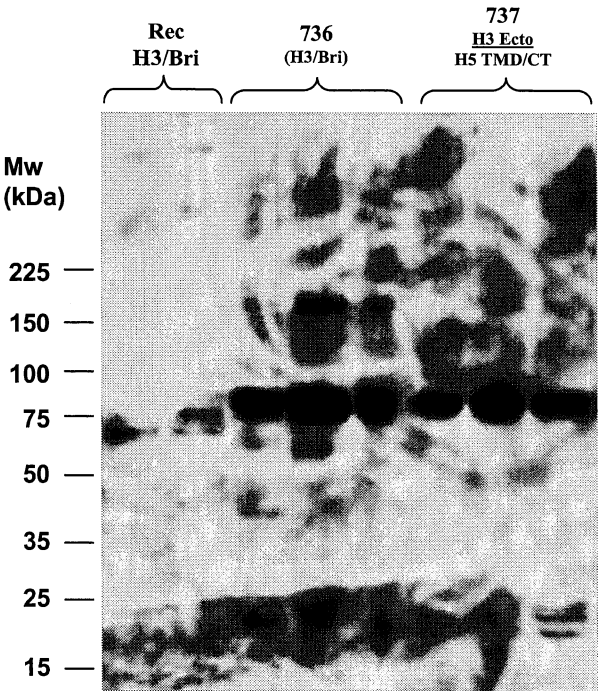
13/49

Fig. 14



Kháng B-Florida

Fig. 15



Kháng H3-Brisbane

Fig. 16

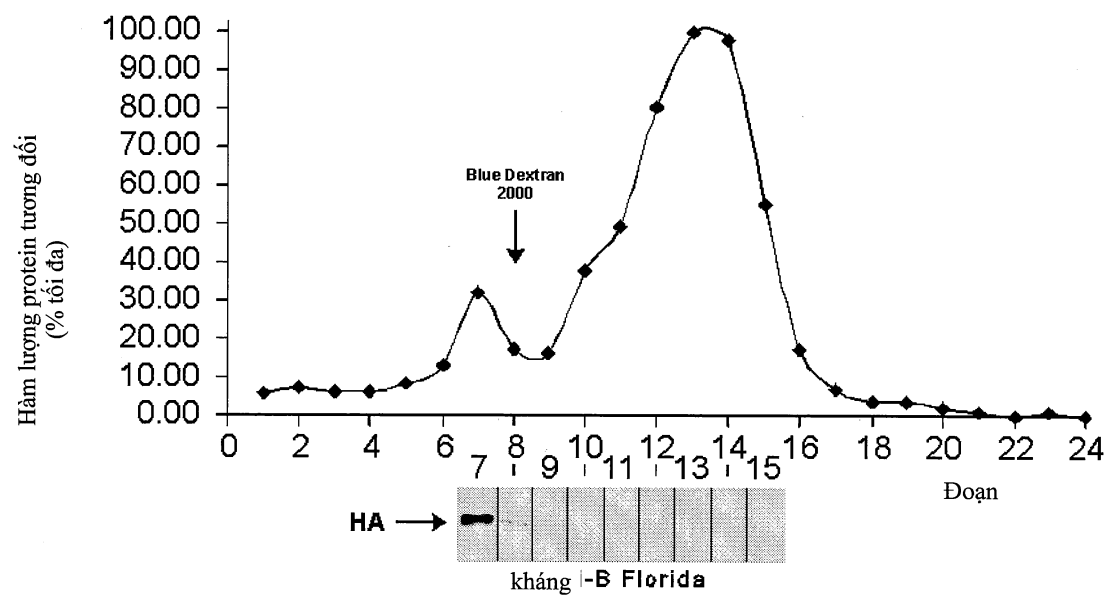


Fig. 17

SEQ. ID. NO. 52

B-Plasto-SP gốc-H5 A/Indonesia/5/05 (HindIII-Sp gốc-H5 A/Indo/5/05-SacI)

AAGCTTATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCAGTCTTGTTAAAAGTGATCAGATT
TGCATTGGTTACCATGCAAACAATTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGT
TACTGTTACACATGCCCCAAGACATACTGGAAAAGACACACAACGGGAAGCTCTGCGATCTAG
ATGGAGTGAAGCCTCTAATTTTAAGAGATTGTAGTGTAGCTGGATGGCTCCTCGGGAACCCA
ATGTGTGACGAATTCATCAATGTACCGGAATGGTCTTACATAGTGGAGAAGGCCAATCCAAC
CAATGACCTCTGTTACCCAGGGAGTTTCAACGACTATGAAGAACTGAAACACCTATTGAGCA
GAATAAACCATTTTGAGAAAATTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCT
CATCAGGAGTTAGCTCAGCATGTCCATACCTGGGAAGTCCCTCCTTTTTAGAAATGTGGTA
TGGCTTATCAAAAAGAACAGTACATACCCAACAATAAAGAAAAGCTACAATAATACCAACCAA
GAGGATCTTTTGGTACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGC
TATATCAAAACCCAACCACCTATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTAC
CAAAAATAGCTACTAGATCCAAAGTAAACGGGCAAAGTGGAAGGATGGAGTTCTTCTGGACA
ATTTTAAAACCTAATGATGCAATCAACTTCGAGAGTAATGGAAATTTCAATTGCTCCAGAATAT
GCATACAAAATTGTCAAGAAAGGGGACTCAGCAATTATGAAAAGTGAATTGGAATATGGTAA
CTGCAACACCAAGTGTCAAACCTCAATGGGGGCGATAAACTCTAGTATGCCATTCCACAACA
TACACCCTCTCACCATCGGGGAATGCCCCAAATATGTGAAATCAAACAGATTAGTCCTTGCA
ACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAAGAAAAAAGAGAGGACTATTTGGAG
CTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGTACCA
CCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTCAAAAGGCAATA
GATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGCCGT
TGGAAGGGAATTTAATAACTTAGAAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACG
GGTTTCTAGATGTCTGGACTTATAATGCCGAACCTCTGGTTCTCATGGAAAATGAGAGAACTC
TAGACTTTTCATGACTCAAATGTTAAGAACCTCTACGACAAGGTCCGACTACAGCTTAGGGAT
AATGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATCACAAATGTGATAATGAATGTATG
GAAAGTATAAGAAACGGAACGTACAACCTATCCGCAGTATTCAGAAGAAGCAAGATTAAGGAG
AGAGGAAATAAGTGGGGTAAAATTGGAATCAATAGGAACTTACCAATACTGTCAATTTATTC
AACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCT
CCAATGGATCGTTACAATGCAGAATTTGCATTTAAGAGCTC

Fig. 18

SEQ. ID. NO. 53

Casset biểu hiện 660 từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu Plastoxyanin) đến EcoRI (ngay cùng chiều với nhân tố kết thúc Plastoxyanin). H5 từ trình tự mã hóa A/Indonesia/5/2055 được gạch dưới.

AAGCTTGCTAGCGGCCTCAATGGCCCTGCAGGTCGACTCTAGAGGTACCCCGGGCTGGTATATTTATA
 TGTGTCAAATAACTCAAAAACCATAAAAGTTTAAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAAT
 ATTCACCTACTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGT
 AGTGATATTTTGACAACAATTTTGTGCAACATTTGAGAAAATTTTGTGTTCTCTCTTTTCATTGGTCAA
 AAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAAGTTGTA
 CAAAAGTTGTACCAAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTT
 AATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTA
 AAAAGAAAGAATAAATTATTTTAAAATTTAAAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAAT
 TGATGAAAGAGTTGGATTAAGTTGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATC
 TTTTCTATATATTGCCCCATAGAGTCAGTTAACTCATTTTATATTTTCATAGATCAAATAAGAGAAATAA
 CGGTATATTAATCCCTCCAAAAAACCGGTATATTTACTAAAAATCTAAGCCACGTAGGAGGATA
 ACAGGATCCCCGTAGGAGGATAACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTT
 TAAGCCACGCATCTGTGGCACATCTACATTATCTAAATCACACATTCTCCACACATCTGAGCCACACA
 AAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCT
 TCAAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAGAAAATAGTG
CTTCTTCTTGCAATAGTCAGTCTTGTTAAAGTGATCAGATTTGCATTGGTTACCATGCAAACAATTCAA
CAGAGCAGGTTGACACAATCATGGAAGAAGCTTACTGTTACACATGCCCAAGACATACTGGAAGA
CACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTAAGAGATTGTAGTGTAGCTG
GATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCGGAATGGTCTTACATAGTGGAG
AAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGAGTTTCAACGACTATGAAGAAGTGAACACCTA
TTGAGCAGAATAAACCATTTTGAGAAAATTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGAAGCCT
CATCAGGAGTTAGCTCAGCATGTCCATACCTGGGAAGTCCCTCCTTTTTAGAAATGTGGTATGGCTTA
TCAAAAAGAACAGTACATACCCAACAATAAGAAAAGCTACAATAATACCAACCAAGAGGATCTTTGGT
ACTGTGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAGGCTATATCAAAACCAACCACT
ATATTTCCATTGGGACATCAACACTAAACCAGAGATTGGTACCAAAAATAGCTACTAGATCCAAAGTAAA
CGGGCAAAGTGAAGGATGGAGTTCTTCTGGACAATTTTAAACCTAATGATGCAATCAACTTCGAGAG
TAATGGAAATTTTATTGCTCCAGAATATGCATACAAAATTGTCAAGAAAGGGGACTCAGCAATTATGAAA
AGTGAATTGGAATATGGTAACTGCAACACCAAGTGTCAAACTCCAATGGGGGCGATAAACTCTAGTATG
CCATTCCACAACATACACCCTCTCACCATCGGGGAATGCCCAAATATGTGAAATCAAACAGATTAGTC
CTTGCAACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAAGAAAAAGAGAGGACTATTTGGAGC
TATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGTACCACCATAGCA
ATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCAATCAAAAGGCAATAGATGGAGTCACCAAT
AAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGCCGTTGGAAGGGAATTTAATAACTTA
GAAAGGAGAAATAGAGAATTTAACAAGAAGATGGAAGACGGGTTTCTAGATGTCTGGACTTATAATGCC
GAACTTCTGGTTCTCATGGAAAATGAGAGAACTCTAGACTTTTATGACTCAAATGTTAAGAACCTCTACG
ACAAGGTCCGACTACAGCTTAGGGATAATGCAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATCACA
AATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAACCTATCCGCAGTATTCAGAAGAAG
CAAGATTAAAAAGAGAGGAAATAAGTGGGGTAAATTTGGAATCAATAGGAACCTTACCAAACTGTCAAT
TTATTCAACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTC
CAATGGATCGTTACAATGCAGAATTTGCATTTAAGAGCTCTAAGTTAAATGCTTCTCGTCTCCTATTTA
 TAATATGGTTTGTATTGTTAATTTTGTCTTGTAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTA
 TTTGTATGAGATGAACGGTGAATGTAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCC
 ATAATACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATTTGAACAACATAAAATTGAACATCT
 TTTGCCACAACCTTTAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAAATGGA
 TATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAACCTACTAATTTTATATCATCCCCTTTGAT
 AAATGATAGTACCAATTAGGAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Fig. 19

SEQ. ID. NO. 54

Trình tự mã hóa H1/NC (H1N1) thiếu TmD và Ctail (Ngân hàng gen số gia nhập AY289929)

ATGAAAGCAAACTACTGGTCCTGTTATGTACATTTACAGCTACATATGCAGACACAATATGTAT
 AGGCTACCATGCCAACAACTCAACCGACACTGTTGACACAGTACTTGAGAAGAATGTGACAGT
 GACACACTCTGTCAACCTACTTGAGGACAGTCACAATGGAAAACTATGTCTACTAAAAGGAATA
 GCCCCACTACAATTGGGTAATTGCAGCGTTGCCGGATGGATCTTAGGAAACCCAGAATGCGAA
 TTAAGTATTTCCAAGGAATCATGGTCCTACATTGTAGAAACACCAAATCCTGAGAATGGAACAT
 GTTACCCAGGGTATTTTCGCCGACTATGAGGAACTGAGGGAGCAATTGAGTTCAGTATCTTCAT
 TTGAGAGATTGCAAATATTTCCCAAGAAAGCTCATGGCCCAACACACCGTAACCGGAGTAT
 CAGCATCATGCTCCCATAATGGGAAAAGCAGTTTTTTACAGAAATTTGCTATGGCTGACGGGGA
 AGAATGGTTTTGTACCCAAACCTGAGCAAGTCCTATGTAAACAACAAAGAGAGAAAGAAGTCCTTGT
 ACTATGGGGTGTTTCATCACCCGCCTAACATAGGGAACCAAAGGGCCCTCTATCATACAGAAAA
 TGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAAGATTCACCCCAGAAATAGCCAAAAGA
 CCCAAAGTAAGAGATCAGGAAGGAAGAATCACTACTACTGGACTCTGCTGGAACCTGGGGAT
 ACAATAATATTTGAGGCAAATGGAAATCTAATAGCGCCATGGTATGCTTTTGCAGTGAAGTAGAG
 GCTTTGGATCAGGAATCATCACCTCAAATGCACCAATGGATGAATGTGATGCGAAGTGTCAAA
 CACCTCAGGGAGCTATAAACAGCAGTCTTCCTTTCCAGAATGTACACCCAGTCACAATAGGAG
 AGTGTCCAAAGTATGTCAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCAT
 CCATTCAATCCAGAGGTTTTGTTTGGAGCCATTGCCGGTTTCATTGAAGGGGGGTGGACTGGAA
 TGGTAGATGGGTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATC
 AAAAAAGTACACAAAATGCCATTAACGGGATTACAAACAAGGTGAATTCTGTAATTGAGAAAAT
 GAACACTCAATTCACAGCTGTGGGCAAAGAATTCAACAAATTGGAAAGAAGGATGGAAAACTTA
 AATAAAAAAGTTGATGATGGGTTTCTAGACATTTGGACATATAATGCAGAATTGTTGGTTCTACT
 GGAAAATGAAAGGACTTTGGATTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAGTAAAA
 AGCCAATTAAGAATAATGCCAAAGAAATAGGAAACGGGTGTTTTGAATTCTATCACAAGTGTA
 ACAATGAATGCATGGAGAGTGTGAAAAATGGAACCTTATGACTATCCAAAATATTCCGAAGAATC
 AAAGTTAAACAGGGAGAAAATTGATGGAGTGAAATGGAATCAATGGGAGTCTATCAGATTCTG
 GCGATCTACTCACTGTCGCCAGTTCCTGGTTCTTTTGGTCTCCCTGGGGGCAATCAGCTTC
 TGGATGTGTTCCAATGGGTCTTGCAGTGTAGAATATGCATCTGAGACCAGAATTTCA

Fig. 20

SEQ. ID. NO. 55

BgIII-H1 A/NC/20/99-SacI/StuI được tổng hợp.

AGATCTTCGCTGACACAATATGTATAGGCTACCATGCCAACAACTCAACCGACACTGTTGACAC
AGTACTTGAGAAGAATGTGACAGTGACACACTCTGTCAACCTACTTGAGGACAGTCACAATGGA
AAACTATGTCTACTAAAAGGAATAGCCCCACTACAATTGGGTAATTGCAGCGTTGCCGGATGGA
TCTTAGGAAACCCAGAATGCGAATTACTGATTTCCAAGGAATCATGGTCCTACATTGTAGAAACA
CCAAATCCTGAGAATGGAACATGTTACCCAGGGTATTTGCGCCGACTATGAGGAACTGAGGGAGC
AATTGAGTTCAGTATCTTCATTTGAGAGATTCGAAATATTTCCCAAAGAAAGCTCATGGCCCAAC
CACACCGTAACCGGAGTATCAGCATCATGCTCCCATAAATGGGAAAAGCAGTTTTTACAGAAATTT
GCTATGGCTGACGGGGAAGAATGGTTTGTACCCAAACCTGAGCAAGTCCTATGTAAACAACAAA
GAGAAAGAAGTCCTTGACTATGGGGTGTTCATCACCCGCCTAACATAGGGAACCAAAGGGCAC
TCTATCATACAGAAAATGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAAGATTCACCCCA
GAAATAGCCAAAAGACCCAAAGTAAGAGATCAGGAAGGAAAGAATCAACTACTACTGGACTCTGC
TGGAACCTGGGGATACAATAATATTTGAGGCAAATGGAAATCTAATAGCGCCATGGTATGCTTTT
GCACTGAGTAGAGGCTTTGGATCAGGAATCATCACCTCAAATGCACCAATGGATGAATGTGATG
CGAAGTGTCAAACACCTCAGGGAGCTATAAACAGCAGTCTTCCTTTCCAGAATGTACACCCAGT
CACAATAGGAGAGTGTCCAAAGTATGTCAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGG
AACATCCCATCCATTCAATCCAGAGGTTTGTGGAGCCATTGCCGGTTTCATTGAAGGGGGGT
GGACTGGAATGGTAGATGGGTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGC
TGCAGATCAAAAAAGTACACAAAATGCCATTAACGGGATTACAAACAAGGTCAATTCTGTAATTG
AGAAAATGAACACTCAATTCACAGCTGTGGGCAAAGAGTTCAACAAATTGGAAGAAGGATGGA
AACTTAAATAAAAAAGTTGATGATGGGTTTCTAGACATTTGGACATATAATGCAGAATTGTTGGT
TCTACTGGAAAATGAAAGGACTTTGGATTTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAG
TAAAAAGCCAATTAAAGAATAATGCCAAAGAAATAGGAAACGGGTGTTTTGAGTTCTATCACAAG
TGTAACAATGAATGCATGGAGAGTGTGAAAAATGGTACCTATGACTATCCAAAATATTCCGAAGA
ATCAAAGTTAAACAGGGAGAAAAATTGATGGAGTGAAATTGGAATCAATGGGAGTATACTAAGAG
CTCAGGCCT

Fig. 21

SEQ. ID. NO. 53

KpnI-H1 A/NC/20/99 TmD+Ctail-SacI/StuI được tổng hợp.

GGTACCTATGACTATCCAAAATATTCCGAAGAATCAAAGTTAAACAGGGAGAAAAATTGATGGAGTG
 AAATTGGAATCAATGGGAGTATACCAGATTCTGGCGATCTACTCAACTGTCGCCAGTTCCCTGGT
 TCTTTTGGTCTCCCTGGGGGCAATCAGCTTCTGGATGTGTTCCAATGGGTCTTTCAGTGTAGAA
 TATGCATCTAAGAGCTCAGGCCT

Fig. 22

SEQ. ID. NO. 57

Peptit báo hiệu của gen protein disulfua isomeraza (PDI) ở cò linh lăng (các nucleotit 21-103; Số gia nhập Z11499).

CCAAATCCTTAACATTCTTTCAACACCAACAATGGCGAAAAACGTTGCGATTTTCGGTTTATTGTTT
 TCTCTTCTTCTGTTGGTTCCTTCTCAGATCTTCGCTGAGGAATCATCAACTGACGCTAAGGAATTT
 GTTCTTACATTGGATAACACTAATTTCCATGACACTGTAAAGAAGCACGATTTTCATCGTCGTTGAAT
 TCTACGCACCTTGGTGTGGACACTGTAAGAAGCTAGCCCCAGAGTATGAGAAGGCTGCTTCTATC
 TTGAGCACTCACGAGCCACCAGTTGTTTTGGCTAAAGTTGATGCCAATGAGGAGCACAAACAAAGA
 CCTCGCATCGGAAAAATGATGTTAAGGGATTCCCAACCATTAAAGATTTTTAGGAATGGTGGAAAGAA
 CATTCAAGAATACAAAGGTCCCCGTGAAGCTGAAGGTATTGTTGAGTATTTGAAAAACAAAGTGG
 CCCTGCATCCACAGAAATTAATCTGCTGATGATGCGACCGCTTTTGTGGTGACAACAAAGTTGT
 TATTGTCGGAGTTTTCCCTAAATTTTCTGGTGAGGAGTACGATAACTTCATTGCATTAGCAGAGAA
 GTTGCGTTCTGACTATGACTTTGCTCACACTTTGAATGCCAAACACCTTCCAAAGGGAGACTCATC
 AGTGTCTGGGCCTGTGGTTAGGTTATTTAAGCCATTTGACGAGCTCTTTGTTGACTCAAAGGATTT
 CAATGTAGAAGCTCTAGAGAAATTCATTGAAGAATCCAGTACCCCAATTGTGACTGTCTTCAACAA
 TGAGCCTAGCAATCACCTTTTGTGTCAAATCTTTAACTCTCCCAACGCAAAGGCTATGTTGTT
 CATCAACTTTTACTACCGAAGGTGCTGAATCTTTCAAACAAAATACCATGAAGTGGCTGAGCAATA
 CAAACAACAGGGAGTTAGCTTTCTTGTGGAGATGTTGAGTCTAGTCAAGGTGCCTTCCAGTATTT
 TGGACTGAAGGAAGAACAAGTACCTCTAATTATTATTACGCATAATGATGGCAAGAAGTTTTTCAA
 ACCCAATTTGGAACCTTGATCAACTCCCAACTTGTTGAAGGCATACAAGGATGGCAAGGTTGAAC
 CATTTGTCAAGTCTGAACCTATTCCTGAACTAACAACGAGCCTGTTAAAGTGGTGGTTGGGCAAA
 CTCTTGAGGACGTTGTTTTCAAGTCTGGGAAGAATGTTTTGATAGAGTTTTATGCTCCTTGGTGTG
 GTCAGTGAAGCAGTTGGCTCCAATCTTGGATGAAGTTGCTGTCTCATTCCAAAGCGATGCTGAT
 GTTGTTATTGCAAACTGGATGCAACTGCCAACGATATCCCAACCGACACCTTTGATGTCCAAGG
 CTATCCAACCTTGTAATTCAGGTCAGCAAGTGGAAACTATCACAATACGACGGTGGTAGGACAA
 AGGAAGACATCATAGAATTCATTGAAAAGAACAAGGATAAACTGGTGCTGCTCATCAAGAAGTAG
 AACAAACAAAAGCTGCTGCTCAGCCAGAAGCAGAACAACCAAAAGATGAGCTTTGAAAAGTTCCG
 CTTGGAGGATATCGGCACACAGTCATCTGCGGGCTTTACAACCTTTTTGTATCTCAGAATCAGAA
 GTTAGGAAATCTTAGTGCCAATCTATCTATTTTTGCGTTTCATTTTATCTTTTTGGTTTACTCTAATG
 TATTACTGAATAATGTGAGTTTTGGCGGAGTTTAGTACTGGAACCTTTGTTTCTGTAAAAA
 AA

Fig. 23A

SEQ. ID. NO. 58

PromPlasto-PDI SP-BgIII

CTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAAAAGTTTAAGTTAGCAAGTGTGTACATT
 TTTACTTTGAACAAAAATATTCACCTACTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATA
 TGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTTGCAACATTTGAGAAAATTTTG
 TTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGAGGGGAGAATAAAAAACATA
 ATGTGAGTATGAGAGAGAAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCATTGAGGA
 ATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGA
 TGTACCATTAGAGAATTTTGGCAAGTCATTAAGAAAGAAAGAATAAATTATTTTAAAAATTTAAAGTT
 GAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGT
 AATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCCATAGAGTCAG
 TTAACTCATTTTATATTTTCATAGATCAAATAAGAGAAAATAACGGTATATTAATCCCTCCAAAAAAA
 AAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATA
 ACATCCAATCCAACCAATCACACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGG
 CACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATC
 TTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACA
 AAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAAATGGCGAAAAACGTTGCGATTTTCG
GCTTATTGTTTTCTCTTCTTGTGTTGGTTCCCTTCTCAGATCT

Fig. 23B

SEQ. ID. NO. 85

Trình tự plasmit SacI-Plasto 3'UTR

GAGCTCTAAGTTAAAATGCTTCTTCGTCTCCTATTTATAATATGGTTTGTTATTGTTAATTTTGTTCT
 TGTAGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAAGTGGTGTAAATG
 TAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCCATAACTAAGTACATGAAGACC
 TGCCGCGTACAATTGCTTATATTTGAACAATAAAATTGAACATCTTTGCCACAACCTTTATAAGT
 GGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCGAAAT
 CATTACAATCAACTTAACGTTATTAATACTAATTTTATATCATCCCCTTTGATAAATGATAGTACA

Fig. 24

SEQ. ID. NO. 59

Casset biểu hiện 540 từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu Plastoxyanin) đến EcoRI (ngay cùng chiều với nhân tố kết thúc Plastoxyanin). H1 từ trình tự mã hóa A/New Caledonia/20/1999 được gạch dưới.

AAGCTTGCTAGCGGCCCTCAATGGCCCTGCAGGTCGACTCTAGAGGTACCCCGGGCTGGTATATTTATATGTTGT
 CAAATAACTCAAAAACCATAAAAGTTTAAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATATTCACCTACTA
 CTGTTATAAATCATTATTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAAC
 AATTTTGTGCAACATTTGAGAAAATTTTGTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGG
 AAGAGGGGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAA
 ATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGA
 GAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAGAAAGAAAGAAATAAATTATTTTAAAATTTAAAGTTGAGTC
 ATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAATTAGAATTTGGT
 GTCAAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCATAGAGTCAGTTAACTCATTTTTATATTTTCATAG
 ATCAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAGAAACGGTATATTTACTAAAAATCTAAGCCA
 CGTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATCCAATCCAACCAATCACACAATCCTGATGAGATAA
 CCCACTTTAAGCCCACGCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACA
 CAAAAACCAATCCACATCTTTATCACCCATTCTAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAA
 ACACATACAAAGAGAAGAGAGACTAATTAATTAATCAATCATCTTGAGAGAAAATGGCGAAAAACGTTGCCGATTTTCG
 GCTTATTGTTTTCTCTTCTGTGTTGGTTCCTTCTCAGATCTTCGCTGACACAATATGTATAGGCTACCATGCCAA
 CAACTCAACCGACACTGTTGACACAGTACTTGAGAAGAATGTGACAGTGACACACTCTGTCAACCTACTTGAGG
 ACAGTCACAATGGAAAACATGTCTACTAAAAGGAATAGCCCCACTACAATTGGGTAATTGCAGCGTTGCCGGAT
 GGATCTTAGGAAACCCAGAATGCGAATTACTGATTTCCAAGGAATCATGGTCCTACATTGTAGAAACACCAAATC
 CTGAGAATGGAACATGTTACCCAGGGTATTTCCGCCACTATGAGGAACTGAGGGAGCAATTGAGTTTCAGTATCT
 TCATTTGAGAGATTGAAAATATTTCCCAAAGAAAGCTCATGGCCCAACCAACCGTAACCGGAGTATCAGCATCA
 TGCTCCCATATGGGAAAAGCAGTTTTTACAGAAATTTGCTATGGCTGACGGGGAAGAATGGTTTGTACCCAAAC
 CTGAGCAAGTCCTATGTAACAACAAAGAGAAAAGAGTCTTGTACTATGGGGTGTTCATCACCCGCCTAACATA
 GGGAAACCAAAGGGCACTCTATCATACAGAAAATGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAAGATTC
 ACCCCAGAAATAGCCAAAAGACCCAAAGTAAGAGATCAGGAAGGAAGAAATCAACTACTACTGGACTCTGCTGGA
 ACCTGGGGATACAATAATTTGAGGCAATGGAAATCTAATAGCGCCATGGTATGCTTTTGCAGTCTGAGTAGAGG
 CTTTGGATCAGGAATCATCACCTCAAATGCACCAATGGATGAATGTGATGCGAAGTGCAAAACACCTCAGGGAG
 CTATAAACAGCAGTCTTCTTTCCAGAATGTACACCCAGTCACAATAGGAGAGTGTCCAAAGTATGTCAGGAGTG
 CAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCATCCATTCAATCCAGAGGTTTGTGTTGGAGCCATTGCCG
 GTTTTCATTGAAGGGGGGTGGACTGGAATGGTAGATGGGTGGTATGGTTATCATCATCAGAATGAGCAAGGATCT
 GGCTATGCTGCAGATCAAAAAGTACACAAAATGCCATTAACGGGATTACAAACAAGGTCAATTCTGTAATTGAG
 AAAATGAACACTCAATTCACAGCTGTGGGCAAGAGTTCAACAAATTGGAAGAAGGATGGAAAACCTAAATAAA
 AAAGTTGATGATGGGTTTCTAGACATTTGGACATATAATGCAGAATTGTTGGTTCTACTGGAAAATGAAAGGACTT
 TGGATTTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAATTAAAGAATAATGCCAAAGAAAT
 AGGAAACGGGTGTTTTGAGTTCTATCACAAAGTGAACAATGAATGCATGGAGAGTGTAAGAAATGGTACCTATGA
 CTATCCAAAATATTCGAAGAATCAAAGTTAAACAGGGAGAAAATTGATGGAGTGAAATTGGAATCAATGGGAGT
 ATACCAGATTCTGGCGATCTACTCAACTGTGCGCCAGTTCCCTGGTCTTTTGGTCTCCCTGGGGGCAATCAGCTT
 CTGGATGTGTTCCAATGGGTCTTTGCAGTGTAGAATATGCATCTAAGAGCTCTAAGTTAAATGCTTCTTCGTCT
 CCTATTTATAATATGGTTTGTATTGTTAATTTTGTCTTGTAAGAGCTTAATTAATCGTTGTTGTTATGAAATAC
 TATTTGTATGAGATGAAGTGGTGAATGTAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCCATAA
 CTAAGTACATGAAGACCTGCCGCTACAATTGTCTTATATTTGAACAACCTAAAATTGAACATCTTTTGCCACAA
 CTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCGAAAT
 TCATTAACAATCAACTTAACGTTATTAATACTAATTTTATATCATCCCCTTGATAAATGATAGTACACCAATTAG
 GAAGGAGCATGCTCGAGGCCCTGGCTGGCCGAATTC

SEQ. ID. NO. 60

CACCTTTGTGAGTCTACACTTTTGATTCCCTTCAACACATACAAAGAGAAGAGACTAATTAATTAAT
TAATCATCTTGAGAGAAAATGAAAGTAAAACTACTGGTCCTGTTATGCACATTTACAGCTACATAT
GCAGACACAATATGTATAGGCTACCATGCTAACAACCTCGACCGACACTGTTGACACAGTACTTGA
AAAGAATGTGACAGTGACACACTCTGTCAACCTGCTTGAGAACAGTCACAATGGA AAACTATGTC
TATTA AAAAGGAATAGCCCCACTACAATTGGGTAATTGCAGCGTTGCCGGGTGGATCTTAGGAAAC
CCAGAATGCGAATTACTGATTTCCAAGGAGTCATGGTCCTACATTGTAGAAAAACCAATCCTGA
GAATGGAACATGTTACCCAGGGCATTTCGCTGACTATGAGGAACTGAGGGAGCAATTGAGTTCA
GTATCTTCATTTGAGAGGTTTCGAAATATTCCCCAAAGAAAGCTCATGGCCCAACCCACACCGTAAC
CGGAGTGTGAGCATCATGCTCCCATAAATGGGGAAAGCAGTTTTTACAGAAATTTGCTATGGCTGA
CGGGGAAGAATGGTTTGTACCCAAACCTGAGCAAGTCCTATGCAAAACAACAAAGAAAAAGAAAGT
CCTTGTACTATGGGGTGTTCTATCCCGCCCAACATAGGTGACCAAAAGGCCCTCTATCATACA
GAAATGCTTATGTCTCTGTAGTGTCTTCCACATTATAGCAGAAATTCACCCGAAATAGCCCAA
AGACCCAAAGTAAGAGATCAAGAAGGAAGTAATCAATTACTACTGGACTCTGCTGAACCCGGGG
ATACAATAATATTTGAGGCAAATGGAATCTAATAGCGCCAAGATATGCTTTGCACTGAGTAGA
GGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAAATGTGATGCGAAGTGCCAAAC
ACCTCAGGGAGCTATAAACAGCAGTCTTCTTTCCAGAACGTACCCAGTCACAATAGGAGAG
TGTCCAAAGTATGTCAGGAGTGCAAAATTAAGGATGTTACAGGACTAAGGAACATCCCATCCAT
TCAATCCAGAGGTTTGTGGAGCCATTGCCGTTTCATTGAAGGGGGGTGGACTGGAATGGTA
GATGGTTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATCAAAAAAG
CACACAAAATGCCATTAATGGGATTACAAACAAGGTCAATTCTGTAATTGAGAAAATGAACACTCA
ATTCACAGCAGTGGGCAAAGAGTTCAACAAATTGGAAGAAGGATGGAAAAC TTGAATAAAAAAG
TTGATGATGGGTTTATAGACATTTGGACATATAATGCAGAACTGTTGGTTCTACTGGAAAATGAAA
GGACTTTGGATTTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAGTTAAAG
AATAAGTCTAAAGAAATAGGAAATGGGTGTTTTGAGTTCTATCACAAGTGTAACGATGAATGCAT
GGAGAGTGTAAGAAGTGAACCTTATGACTATCCAAAATATCCGAAGAACTCAAAGTTAAACAGGG
AGAAAATTGATGGAGTGAAATTGGAATCAATGGGAGTCTATCAGATTCTGGCGATCTACTCAACA
GTCGCCAGTTCTCTGGTTCTTTTGGTCTCCCTGGGGCAATCAGCTTCTGGATGTGTTCCAATG
GGTCTTTACAGTGTAGAATATGCATCTAAGAGCTC

Fig. 26

SEQ. ID. NO. 61

Casset biểu hiện 774 từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu Plastoxyanin) đến EcoRI (ngay cùng chiều với nhân tố kết thúc Plastoxyanin). H1 từ trình tự mã hóa A/Brisbane/59/2007 được gạch dưới.

AAGCTTGCTAGCGGCCCTCAATGGCCCTGCAGGTGCACTCTAGAGGTACCCCGGGCTGGTATATTTATATGTT
GTCAAATAACTCAAAAACCATAAAAGTTTAAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAATATTCACC
TACTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTT
GACAACAATTTTGTGCAACATTTGAGAAAATTTGTTGTTCTCTCTTTTCATTGGTCAAAAACATAGAGAGA
GAAAAAGGAAGAGGGGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAAGTTGTACAAAAGTTGTACCAAAA
TAGTTGTACAAAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGG
ATGACGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAAAAGAAAGAATAAATTTATTTTAA
AATTAAGTTGAGTCATTTGATTAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTAT
TAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCATAGAGTCAGTTA
ACTCATTTTTATATTTATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAATAACGGTAT
ATTTACTAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATCCAATCCAACCA
ATCACAACAATCCTGATGAGATAACCCACTTTAAGCCACGCATCTGTGGCACATCTACATTATCTAAATCAC
ACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACT
TTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGA
GAGAAAATGAAAGTAAACTACTGGTCCTGTTATGCACATTTACAGCTACATATGCAGACACAATATGTATAG
GCTACCATGCTAACAACCTCGACCGACACTGTTGACACAGTACTTGAAAAGAATGTGACAGTGACACACTCTG
TCAACCTGCTTGAGAACAGTCACAATGGAAAATATGTCTATTTAAAGGAATAGCCCCACTACAATTGGGTAA
TTGCAGCGTTGCCGGGTGGATCTTAGGAAACCCAGAATGCCAATTACTGATTTCCAAGGAGTCATGGTCCTA
CATTGTAGAAAAACCAATCCTGAGAATGGAACATGTTACCCAGGGCATTTCGCTGACTATGAGGAACTGAG
GGAGCAATTGAGTTCAGTATCTTCATTTGAGAGGTTGGAATATTTCCCAAAAGAAAGCTCATGGCCCAACCA
CACCGTAACCGGAGTGTGAGCATCATGCTCCCATAAATGGGGAAAGCAGTTTTTACAGAAATTTGCTATGGCT
GACGGGGAAGAAATGGTTTGTACCCAAACCTGAGCAAGTCCTATGCAAACACAAAGAAAAAGAAGTCCTTGT
ACTATGGGGTGTTCATCACCCGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAAATGCTTATGT
CTCTGTAGTGTCTTCAACTTATAGCAGAAAATTCACCCACGAAATAGCCAAAAGACCCAAAGTAAGAGATCAA
GAAGGAAGAATCAATTACTACTGGACTCTGCTTGAACCCGGGGATACAATAATATTTGAGGCAATGGAAAT
CTAATAGCGCCAAGATATGCTTTCGCACTGAGTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCA
ATGGATAAATGTGATGCCAAGTGCCAAACACCTCAGGGAGCTATAAACAGCAGTCTTCCTTTCCAGAACGTA
CACCCAGTCACAATAGGAGAGTGTCCAAAGTATGTCAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGG
AACATCCCATCCATTCAATCCAGAGGTTTGTGGAGCCATTGCCGGTTTCATTGAAGGGGGTGGACTGGA
ATGGTAGATGGTTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATCAAAAAGC
ACACAAAATGCCATTAATGGGATTACAAACAAGGTCAATTCTGTAATTGAGAAAATGAACACTCAATTCACAG
CAGTGGGCAAGAGTTCAACAAATTGGAAAGAAGGATGGAAAACCTGAATAAAAAAGTTGATGATGGGTTTAT
AGACATTTGGACATATAATGCAGAACTGTTGGTTCTACTGGAAAATGAAAGGACTTTGGATTTCCATGACTCC
AATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAGTTAAAGAATAATGCTAAAGAAATAGGAAATGGGTGTT
TTGAGTTCTATCACAAGTGAACGATGAATGCATGGAGAGTGTAAAGAATGGAACCTTATGACTATCCAAAATA
TTCCGAAGAATCAAAGTTAAACAGGGAGAAAATTGATGGAGTGAAATTGGAATCAATGGGAGTCTATCAGATT
CTGGCGATCTACTCAACAGTCGCCAGTTCTCTGGTTCTTTTGGTCTCCCTGGGGGCAATCAGCTTCTGGATG
TGTTCCAATGGGTCTTTACAGTGTAGAATATGCATCTAAGAGCTCTAAGTTAAATGCTTCTCTGCTCCTATT
TATAATATGGTTTGTATTGTTAATTTTGTCTTGTAGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATT
TGTATGAGATGAAGTGGTGAATGTAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCCTCCATAACT
AACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATTTGAACAACTAAAATTGAACATCTTTTGCCACAA
CTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCGA
AATTCATTAACAATCAACTTAACGTTATTAACACTAATTTTATATCATCCCCTTTGATAAATGATAGTACACCA
ATTAGGAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Fig. 27

SEQ. ID. NO. 62

Casset biểu hiện số 828, từ PacI (ngược chiều với trình tự khởi đầu) đến AscI (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự CPMV HT 3'UTR được gạch dưới với ATG đột biến được in đậm. Vị trí giới hạn Apal (ngay cùng chiều ATG với trình tự mã hóa protein, trong trường hợp này chuỗi nhẹ C5-1 kappa, được gạch dưới in nghiêng).

TTAATTAAGAATTGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTT
TATTGAGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGC
CATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCG
TGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGT
AAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCCTCCTCTATATAAGGAAGTTCATTTTCAT
TTGGAGAGGTATTAATAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCT
TCTTCTAAACTCTCTCATCTCTCTTAAGCAAACCTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTT
CAACGTTGTGAGATCGTGCTTCGGCACCAAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATC
TCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTAAGTGTCTTGTCTGCGGTGTGGTC
TTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTGACGCCCCATACATTACTTGTACGATTCTGCTGA
CTTTGCGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTT
CTTCTTGCTGATTGGTTCTATAAGAAATAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCG
AAGTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGCCCATGGTTTTCA
CCTCAGATACTTGGACTTATGCTTTTTTGGATTTGAGCCTCCAGAGGTGATATTGTGCTAAGTCAAGT
CTCCAGCCACCCTGTCTGTGACTCCAGGAGATAGTGTGAGTCTTTCTGTCAGGGCCAGCCAAAGTA
TTAGCAACAACCTACACTGGTTTCAACAAAAATCGCATGAGTCTCCAAGGCTTCTCATCAAGTATGC
TTCCAGTCCATATCTGGGATCCCCTCCAGGTTGAGTGGCAGTGGATCTGGGACAGATTTCACTCT
CAGTATCAACAGTGTGAAGACTGAAGATTTTGAATGTTTTCTGTCAACAGAGTAACAGCTGGCCT
CTCACGTTCCGTGATGGGACAAAGCTGGAGCTGAAACGGGCTGATGCTGCACCAACTGTATCCAT
CTTCCCACCATCCAGTGAGCAGTTAACATCTGGAGGTGCCTCAGTCGTGTGCTTCTTGAACAACCTT
CTACCCCAAAGACATCAATGTCAAGTGAAGATTGATGGCAGTGAACGACAAAATGGCGTCCTGAA
CAGTTGGACTGATCAGGACAGCAAAGACAGCACCTACAGCATGAGCAGCACCCCTCACGTTGACCA
AGGACGAGTATGAACGACATAACAGCTATACCTGTGAGGCCACTCACAAGACATCAACTTACCCCA
TTGTCAAGAGCTTCAACAGGAATGAGTGTAGAGGCCATTTTTCTTTAGTTTGAATTTACTGTTATTC
GGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTAAT
TTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTCAGCAAGGACACAAAAAGATTTTAATTTTATTA
AAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATT
TGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTT
GAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGA
TTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAAACAAATATAGCGCGCAAACCTAGGATAA
ATTATCGCGCGCGGTGTCATCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Fig. 28

SEQ. ID. NO. 63

Sản phẩm thiết kế số 690, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu Plastoxyanin) đến EcoRI (ngay cùng chiều với nhân tố kết thúc Plastoxyanin). Trình tự mã hóa HA thể khảm được gạch dưới.

AAGCTTGCTAGCGGCCTCAATGGCCCTGCAGGTCGACTCTAGAGGTACCCCGGGCTGGTATATTTATATGTT
 GTCAAATAACTCAAAAACCATAAAAGTTTAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATATTCACCT
 ACTACTGTTATAAATCATTATTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTG
 ACAACAATTTTGTGCAACATTTGAGAAAATTTGTTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGA
 AAAAGGAAGAGGGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAATAG
 TTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATG
 ACGCATTAGAGAGATGTACCATTAGAGAATTTTTGGCAAGTCATTAAAAAGAAAGAAATAAATTATTTTTAAAATT
 AAAAGTTGAGTCATTTGATTAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGT
 AATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCCATAGAGTCAGTTAACTC
 ATTTTTATATTTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAACAACGGTATATTT
 ACTAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCCAACCAATCA
 CAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACATCTACATTATCTAAATCACACATT
 CTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTG
 AGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAA
 AATGGAGAAAAATAGTGCTTCTTCTGCAATAGTCAGTCTTGTAAAAAGTGATCAGATTTGCATTTGGTTACCATG
 CAAACAATTCACAGAGCAGGTTGACACAATCATGGAAAAAGAACGTTACTGTTACACATGCCCAAGACATACT
 GGAAAAAGACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTTAAGAGATTGTAGTGT
 AGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCGGAATGGTCTTACATAGTGGA
 GAAGGCCAATCCAACCAATGACCTCTGTTACCCAGGGCATTTGCTGACTATGAGGAACTGAGGGAGCAATT
 GAGTTCAGTATCTTCATTTGAGAGGTTGCAAAATATCCCAAGAAAGCTCATGGCCCAACCAACCGTAACC
 GGAGTGTGAGCATCTCCCATAAATGGGGAAGCAGTTTTTACAGAAATTTGCTATGGCTGACGGGGAAG
 AATGGTTGTACCCAAACCTGAGCAAGTCCTATGCAAAACAAGAAAAAGAAAGTCTTGTACTATGGGGTG
 TTCATCACCCGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAAATGCTTATGTCTCTGTAGTGTC
 TTCACATTATAGCAGAAAAATTCACCCAGAAATAGCCAAAAGACCCAAAGTAAGAGATCAAGAAGGAAGAATC
 AATTACTACTGGACTCTGCTTGAACCCGGGGATACAATAATATTTGAGGCAAATGGAATCTAATAGCGCCAA
 GATATGCTTTCGCACTGAGTAGAGGCTTTGGATCAGGAATTATGAAAAGTGAATTGGAATATGGTAACTGCAA
 CACCAAGTGTCAAACCTCAATGGGGGCGATAAACTCTAGTATGCCATTCCACAACATACACCTCTCACCATC
 GGGGAATGCCCAAATATGTGAAATCAAACAGATTAGTCCTTGCAACAGGGCTCAGAAATAGCCCTCAAAGA
 GAGAGCAGAAGAAAAAGAGAGGACTATTTGGAGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAAT
 GGTAGATGGTTGGTATGGGTACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCA
 CTCAAAAGGCAATAGATGGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTGAGGC
 CGTTGGAAGGGAATTTAATAACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACGGGTTTCTA
 GATGTCTGGACTTATAATGCCGAATTTCTGGTTCTCATGGAAAATGAGAGAACTCTAGACTTTTCATGACTCAA
 ATGTTAAGAACCTCTACGACAAGGTCCGACTACAGCTTAGGGATAATGCAAAGGAGCTGGGTAAACGGTTGTT
 TCGAGTTCTATCACAATGTGATAATGAATGTATGGAAGTATAAGAAACGGAACGTACAACATATCCGCAGTA
 TTCAGAAGAAGCAAGATTAAAAAGAGAGGAAATAAGTGGGGTAAAATTGGAATCAATAGGAACCTTACCAAATA
 CTGTCAATTTATTCACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGT
 GCTCCAATGGATCGTTACAATGCAGAATTTGCATTTAAGAGCTCTAAGTTAAATGCTTCTTCGTCTCTATTT
 ATAATATGGTTTGTATTGTTAATTTTGTCTTTAGAGAAGCTTAATTAATCGTTGTTGTTATGAAATACTATTT
 GTATGAGATGAACTGGTGTAATGTAATTCATTACATAAGTGGAGTCAGAATCAGAATGTTTCTCCATAACTA
 ACTAGACATGAAGACCTGCCGCTACAATTGTCTTATATTTGAACAACATAAATTGAACATCTTTTCCACAAC
 TTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGAAATATCAGTTATCGAAA
 TTCATTAACAATCAACTTAACGTTATTAACCTACTAATTTTATCATCCCTTTGATAAATGATAGTACACCAATT
 AGGAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Fig. 29

SEQ. ID. NO. 64

Sản phẩm thiết kế số 691, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu Plastoxyanin) đến EcoRI (ngay cùng chiều với nhân tố kết thúc Plastoxyanin). Trình tự mã hóa HA thể khảm được gạch dưới.

AAGCTTGCTAGCGGCCTCAATGGCCCTGCAGGTCGACTCTAGAGGTACCCCGGGCTGGTATATTTATATGTTGTC
 AAATAACTCAAAAACCATAAAAGTTTAAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTACTACT
 GTTATAAATCATTATTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATT
 TTGTTGCAACATTTGAGAAAATTTGTTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGAG
 GGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAAATATCAT
 TGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGCATTAGAGAGATGT
 ACCATTAGAGAATTTTTGGCAAGTCATTA AAAAGAAAGAATAAATTATTTTTAAAATTTAAAGTTGAGTCATTTGATT
 AAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAGTTGTATTAGTAATTAGAATTTGGTGTCAAATTT
 AATTTGACATTTGATCTTTTCTATATATTGCCCATAGAGTCAGTTAACTCATTTTTATATTTTCATAGATCAAATAAG
 AGAAATAACGGTATATTAATCCCTCCAAAAA AAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGAT
 AACAGGATCCCCGTAGGAGGATAACATCCAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCC
 CACGCATCTGTGGCACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCAC
 ATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAACACATACAAAGAGAA
 GAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAGAAAATAGTGCTTCTTCTTGCAATAGTCAGTCTTGTTA
AAAGTGATCAGATTTGCATTGGTTACCATGCAAACAATTCAACAGAGCAGGTTGACACAATCATGGAAAAGAACGT
TACTGTTACACATGCCAAGACATACTGGAAAAGACACACAACGGGAAGCTCTGCCTATTTAAAGGAATAGCCCC
ACTACAATTGGTAATTGCAGCGTTGCCGGGTGGATCTTAGGAAACCCAGAATGCGAATTACTGATTTCCAAGGA
GTGATGGTCCTACATTGTAGAAAAACCAAAATCCTGAGAATGGAACATGTTACCCAGGGCATTTTCGCTGACTATGAG
GAAGTGAAGGAGCAATTGAGTTGAGTATCTTCATTTGAGAGGTTGGAATATTCCCCAAAGAAAGCTCATGGCCCA
ACCACACCGTAACCGGAGTGTCAGCATCATGCTCCATAATGGGGAAAGCAGTTTTTACAGAAATTTGCTATGGCT
GACGGGGAAGAATGGTTTGTACCCAAACCTGAGCAAGTCCTATGCAAAACAACAAAGAAAAAGAGTCCCTGTACT
ATGGGGGTGTTTCATCACCCGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAAATGCTTATGTCTCTGTA
GTGCTTTCACATTATAGCAGAAAATTCACCCAGAAAATAGCCAAAAGACCCAAAGTAAGAGATCAAGAAGGAAGAA
TCAATTACTACTGGACTCTGCTTGAACCCGGGGATACAATAATTTTTGAGGCAAAATGGAATCTAATAGCGCCAAG
ATATGCTTTGCACTGAGTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAATGCAACACC
AAGTGTCAAACCTCCAATGGGGGCGATAAACTCTAGTATGCCATTCCACAACATACACCCTCTCACCATCGGGGAAT
GCCCCAAATATGTGAAATCAAACAGATTAGTCTTGCACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAA
GAAAAAGAGAGGACTATTTGGAGCTATAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGT
ATGGGTACCACCATAGCAATGAGCAGGGGAGTGGGTACGCTGCAGACAAAGAATCCACTCAAAAGGCAATAGAT
GGAGTCACCAATAAGGTCAACTCAATCATTGACAAAATGAACACTCAGTTTTGAGGCCGTTTGAAGGGGAATTTAATA
ACTTAGAAAGGAGAATAGAGAATTTAAACAAGAAGATGGAAGACGGGTTTTCTAGATGTCTGGACTTATAATGCCGA
ACTTCTGGTTCTCATGGAAAATGAGAGAACTCTAGACTTTTCATGACTCAAATGTTAAGAACCTCTACGACAAGGTC
CGACTACAGCTTAGGGATAATGCAAAGGAGCTGGGTAACGGTTGTTTCGAGTTCTATCACAATGTGATAATGAAT
GTATGGAAAGTATAAGAAACGGAACGTACAACATCCGCAGTATTCAGAAGAAGCAAGATTA AAAAGAGAGGAAAT
AAGTGGGGTAAAATTGGAATCAATAGGAACCTTACCAATACTGTCAATTTATTCAACAGTGCCGAGTTCCCTAGCA
CTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAGAAATTTGCATTTAAG
AGCTCTAAGTTAAAATGCTTCTTCTGCTCTCTATTTATAATATGGTTTGTTATTGTTAATTTGTTCTGTAGAAGAGC
TTAATTAATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAAGTGGTGAATGTAATTCATTTACATAAGTGGAG
TCAGAATCAGAATGTTTCCTCCATAACTAGACATGAAGACCTGCCGCTACAATTGTCTTATATTTGAACAAC
TAAAATTGAACATCTTTTGCCACAACCTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAAT
AATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAATACTACTAATTTTATATCATCCCTTTGA
TAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Fig. 30

SEQ. ID. NO. 65

Sản phẩm thiết kế số 696, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu Plastoxyanin) đến EcoRI (ngay cùng chiều với nhân tố kết thúc Plastoxyanin). Trình tự mã hóa HA thể khảm được gạch dưới.

AAGCTTGCTAGCGGCCTCAATGGCCCTGCAGGTCGACTCTAGAGGTACCCCGGGCTGGTATATTTATATGTTG
TCAAATAACTCAAAAACCATAAAAGTTTAAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATATTCACCTAC
TACTGTTATAAATCATTATTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGAC
AACAATTTTGTGCAACATTTGAGAAAATTTTGTGTTCTCTCTTTTTCATTGGTCAAAAACAATAGAGAGAGAAA
AAGGAAGAGGGAGAATAAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAATAGTT
GTACAAATATCATTGAGGAATTTGACAAAAGCTACACAAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGAC
GCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAGAAAGAAAGAATAAATTTATTTTAAATTA
AAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTATTAGTAAT
TAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCATAGAGTCAGTTAACTCATTT
TTATATTTTATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAGGATATATTTACTAA
AAAATCTAAGCCACGTAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCCAACCAATCACAAACAA
TCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACATCTACATTATCTAAATCACACATTTCTCCA
CACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCTATTCTATAAAAAATCACACTTTGTGAGTCTAC
ACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATCATCTTGAGAGAAAATGGCGA
AAAACGTTGCGATTTTCGGCTTATTGTTTTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCTGACACAATAT
GTATAGGCTACCATGCCAACAACTCAACCGACACTGTTGACACAGTACTTGAGAAGAATGTGACAGTGACACA
CTCTGTCAACCTACTTGAGGACAGTCACAATGGAAAACATGTCTACTAAAAGGAATAGCCCCACTACAATTGG
GTAATTGCAGCGTTGCCGGATGGATCTTAGGAAACCGAATGCGAATTACTGATTTCCAAGGAATCATGGTC
CTACATTGTAGAAACACCAATCCTGAGAATGGAACATGTTACCCAGGGAGTTTCAACGACTATGAAGAACTG
AAACACCTATTGAGCAGAATAAACCATTTTGAAGAAATTCAAATCATCCCCAAAAGTTCTTGGTCCGATCATGA
AGCCTCATCAGGAGTTAGCTCAGCATGTCCATACCTGGGAAGTCCCTCCTTTTTAGAAATGTGGTATGGCTTA
TCAAAAAGAACAGTACATACCCAACAATAAAGAAAAGCTACAATAATACCAACCAAGAGGATCTTTTGGTACTG
TGGGGAATTCACCATCCTAATGATGCGGCAGAGCAGACAAAGGCTATATCAAAACCCAACCATATATTTCCA
TTGGGACATCAACACTAAACCAGAGATTGGTACCAAAAAATAGCTACTAGATCCAAAGTAAACCGGGCAAAGTGG
AAGGATGGAGTTCTTCTGGACAATTTTAAACCTAATGATGCAATCAACTTCGAGAGTAATGGAATTTTCATTG
CTCCAGAATATGCATACAAAATTGTCAAGAAAGGGGACTCAGCAATCATCACCTCAAATGCACCAATGGATGA
ATGTGATGCGAAGTGTCAAACACCTCAGGGAGCTATAAACAGCAGTCTTCTTTCCAGAATGTACACCCAGTC
ACAATAGGAGAGTGTCCAAAGTATGTGAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCAT
CCATTCAATCCAGAGGTTTTGTTTGGAGCCATTGCCGGTTTCATTGAAGGGGGGTGGACTGGAATGGTAGATG
GGTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATCAAAAAAGTACACAAAATGC
CATTAACGGGATTACAAACAAGGTCAATTCTGTAATTGAGAAAATGAACACTCAATTCACAGCTGTGGGCAAAG
AGTTCAACAAATTTGGAAGAAGGATGGAAAACCTAAATAAAAAAGTTGATGATGGGTTTCTAGACATTTGGACA
TATAATGCAGAATTGTTGGTTCTACTGGAAAATGAAAGGACTTTGGATTTCCATGACTCCAATGTGAAGAATCT
GTATGAGAAAAGTAAAAAGCCAAATTAAGAATAATGCCAAGAAATAGGAAACGGGTGTTTTGAGTTCTATCACA
AGTGTAAACAATGAATGCATGGAGAGTGTGAAAAATGGTACCTATGACTATCCAAAATATCCGAAGAATCAAAG
TTAAACAGGGAGAAAATTGATGGAGTGAAATTGGAATCAATGGGAGTATACCAGATTCTGGCGATCTACTCAA
CTGTGCGCCAGTTCCCTGGTTCTTTTGGTCTCCCTGGGGGCAATCAGCTTCTGGATGTGTTCCAATGGGTCTTT
GCAGTGTAGAATATGCATCTAAGAGCTCTAAGTTAAATGCTTCTTCTGCTCCTATTTATAATATGGTTTGTAT
TGTTAATTTTGTCTTGTAGAAGAGCTTAATTAATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAAGTGGT
GTAATGTAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCTCCATAACTAACTAGACATGAAGACCT
GCCGCGTACAATTGTCTTATTTTGAACAACTAAAATTGAACATCTTTTGCCACAACCTTTATAAGTGGTTAATAT
AGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGTTATCGAAATTCATTAAACAATCAACTTA
ACGTTATTAATACTAATTTTATATCATCCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGA
GGCCTGGCTGGCCGAATTC

Fig. 31

SEQ. ID. NO. 66

Sản phẩm thiết kế 732, từ PacI (ngược chiều với trình tự khởi đầu 35S) đến Ascl (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa H1 A/Brisbane/59/2007 được gạch dưới.

TTAATTAAGAATTCGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAA
GATAGTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTC
TGCCGACAGTGGTCCCAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAGAAAGACGTTCCAACCACGT
CTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGA
CCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATACTTAATAGGTTTTGATAAAAGCGAACGT
GGGGAACCCGAACCAAACTTCTTCTAACTCTCTCTCATCTCTCTTAAAGCAAACTTCTCTTTGTCTTTCTTGC
GTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAG
ATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTGGTGTGGTCTTGGGA
AAAGAAAGCTTGCTGGAGGCTGCTGTTGAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCA
ATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTAATCTTTCTTCTTCTTCTTCTGCTGATTGTTCTATAAGAA
ATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGTTTTTTCGAACCTGGAGAAAGATTGTTAAGCTTCTGTATATTCT
GCCCAAATTTGTGGGCCCCATGAAAGTAACTACTGGTCTGTTATGCACATTTACAGCTACATATGCAGACACA
ATATGTATAGGCTACCATGCTAACAACTCGACCGACACTGTTGACACAGTACTTGAAAGAATGTGACAGTGACAC
ACTCTGTCAACCTGCTTGAGAACAGTCACAATGGAAGTATGTCTATTAAGGAATAGCCCCACTACAATTGGG
TAATTGCAGCGTTGCCGGGTGGATCTTAGGAAACCCAGAATGCGAATTACTGATTTCCAAGGAGTCATGGTCTTAC
ATTGTAGAAAAACCAATCCTGAGAATGGAACATGTTACCCAGGGCATTTCGCTGACTATGAGGAAGTGAAGGAG
CAATTGAGTTCAGTATCTTCATTTGAGAGGTTTCGAAATATTCCCAAAGAAAGCTCATGGCCCCAACCAACCGTAA
CCGGAGTGTCAGCATCATGCTCCCATATGGGGAAGCAGTTTTTACAGAAATTTGCTATGGCTGACCGGGGAAGA
ATGGTTTGTACCCAAACCTGAGCAAGTCTATGCAACAACAAAGAAAAAGAGTCTTGTACTATGGGGTGTTCAT
TCACCCGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAATGCTTATGTCTCTGTAGTGTCTTCACATT
ATAGCAGAAATTCACCCAGAAATAGCCAAAAGACCCAAAGTAAGAGATCAAGAAGGAAGAAATCAATTACTACTG
GACTCTGCTGAACCCGGGGATACAATAATTTTGGGCAAATGGAATCTAATAGCGCCAAGATATGCTTTCGCA
CTGAGTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAAATGTGATGCGAAGTGCCAAACAC
CTCAGGGAGCTATAAACAGCAGTCTTCTTTCCAGAACGTACACCCAGTCACAATAGGAGAGTGTCCAAAGTATGT
CAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCATCCATTCAATCCAGAGGTTTGTGAGGCC
ATTGCCGTTTTATTGAAGGGGGGTGGACTGGAATGGTAGATGGTTGGTATGTTATCATCATCAGAATGAGCAA
GGATCTGGCTATGCTGCAGATCAAAAAAGCACACAAAATGCCATTAATGGGATTACAAAACAAAGGTCAATTCTGTAA
TTGAGAAAATGAACACTCAATTCACAGCAGTGGGCAAAGAGTTCAACAAATTGGAAGGAAGGATGGAAGAACTTGAA
TAAAAAGTTGATGATGGGTTTATAGACATTTGGACATATAATGCAGAACTGTTGGTTCTACTGGAAAATGAAAGGA
CTTTGGATTTCCATGACTCCAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAGTTAAAGAATAATGCTAAAGAA
ATAGGAAATGGGTGTTTTGAGTTCTATCACAAGTGTAAACGATGAATGCATGGAGAGTGTAAAGAATGGAACCTATG
ACTATCCAAAATATTCCGAAGAATCAAAGTTAAACAGGGAGAAAATTGATGGAGTGAAATTGGAATCAATGGGAGT
CTATCAGATTCTGGCGATCTACTCAACAGTCGCCAGTCTCTGGTTCTTTGGTCTCCCTGGGGGCAATCAGCTTC
TGGATGTGTTCCAATGGGTCTTTACAGTGTAGAATATGCATCTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTA
TTCGGTGTGCAATTTCTATGTTTGGTGAGCGTTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTAATTTCTTTG
TGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTTATTAATAAAAAAAAAAAAAA
AGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAGTTTCTTAAGATTGAA
TCCTGTTGCCGGTCTTGCAGTATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATG
CATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAAT
ATAGCGCGCAAACTAGGATAAATATCGCGCGCGGTGTCATCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGC
GCGCC

Fig. 32

SEQ. ID. NO. 67

Trình tự mã hóa, từ ATG đến kết thúc, của sản phẩm thiết kế trung gian 787.

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCG
 CTGACACAATATGTATAGGCTACCATGCTAACAACTCGACCGACACTGTTGACACAGTACTTGAAAAAGA
 ATGTGACAGTGACACACTCTGTCAACCTGCTTGAGAACAGTCACAATGGAAAACTATGTCTATTAAGG
 AATAGCCCCACTACAATTGGGTAATTGCAGCGTTGCCGGGTGGATCTTAGGAAACCCAGAATGCGAATT
 ACTGATTTCCAAGGAGTCATGGTCCTACATTGTAGAAAAACCAAATCCTGAGAATGGAACATGTTACCCA
 GGGCATTTCGCTGACTATGAGGAAGTGAAGGAGCAATTGAGTTCAGTATCTTCATTTGAGAGGTTTCGAA
 ATATTCCCCAAAGAAAGCTCATGGCCCAACCAACCGTAACCGGAGTGTCAGCATCATGCTCCATAAT
 GGGGAAAGCAGTTTTTACAGAAATTTGCTATGGCTGACGGGGAAGAATGGTTTGTACCCAAACCTGAGC
 AAGTCCTATGCAACAACAAAGAAAAAGAGTCCTTGACTATGGGGTGTTTCATCACCCGCCAAACATA
 GGTGACCAAAAAGGCCCTCTATCATACAGAAAATGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAA
 AATTCACCCAGAAAATAGCCAAAAGACCCAAAGTAAGAGATCAAGAAGGAAGAATCAATTACTACTGGA
 CTCTGCTTGAACCCGGGGATACAATAATTTTGAAGGCAAATGGAAATCTAATAGCGCCAAGATATGCTTT
 CGCACTGAGTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAAATGTGATGCGAA
 GTGCCAAACACCTCAGGGAGCTATAAACAGCAGTCTTCCTTCCAGAACGTACACCCAGTCACAATAGG
 AGAGTGTCCAAAGTATGTCAGGAGTGCAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCATCCAT
 TCAATCCAGAGGTTTGTGTTGGAGCCATTGCCGGTTTCATTGAAGGGGGGTGGACTGGAATGGTAGATG
 GTTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGGCTATGCTGCAGATCAAAAAGCACACAAA
 ATGCCATTAATGGGATTACAAACAAGGTCAATTCTGTAATTGAGAAAATGAACACTCAATTCACAGCAGT
 GGGCAAAGAGTTCAACAAATTGAAAAGAAGGATGGAAAACCTTGAATAAAAAAGTTGATGATGGGTTTATA
 GACATTTGGACATATAATGCAGAACTGTTGGTTCTACTGGAAAATGAAAGGACTTTGGATTTCCATGACT
 CCAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAGTTAAAGAATAATGCTAAAGAAATAGGAAATGG
 GTGTTTTGAGTTCTATCACAAGTGTAAACGATGAATGCATGGAGAGTGTAAGAATGGAACCTTATGACTAT
 CCAAAATATTCCGAAGAATCAAAGTTAAACAGGGAGAAAATTGATGGAGTGAAATTGGAATCAATGGGA
 GTCTATCAGATTCTGGCGATCTACTCAACAGTCGCCAGTTCTCTGGTTCTTTTGGTCTCCCTGGGGGCA
 ATCAGCTTCTGGATGTGTTCCAATGGGTCTTTACAGTGTAGAATATGCATCTAA

Fig. 33

SEQ. ID. NO. 68

Sản phẩm thiết kế 733, từ PacI (ngược chiều với trình tự khởi đầu 35S) đến Ascl (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa SpPDI A/Brisbane/59/2007 được gạch dưới.

TTAATTAAGAATTGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAAG
 ATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTG
 CCGACAGTGGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTT
 CAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCCT
 TCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATAGGTTTTGATAAAAGCGAACGTGGGGA
 AACCCGAACCAAACCTTCTCTAACTCTCTCATCTCTCTTAAGCAAACCTTCTCTTGTCTTTCTTGCCTGAGC
 GATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTT
 GTGGACACGTAGTGCAGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGGAAAAAGAAAGC
 TTGCTGGAGGCTGCTGTTGAGCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGAATATCTCTAC
 TTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAGTATT
 TCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACCTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGT
 CGGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGC
 TGACACAATATGTATAGGCTACCATGCTAACAACCTGACCGACACTGTTGACACAGTACTTGAAAAGAATGTGACAG
 TGACACACTCTGTCAACCTGCTTGAGAACAGTCACAATGGAAAACCTATGCTATTAAAAGGAATAGCCCCACTACAAT
 TGGGTAATTGCAGCGTTGCCGGGTGGATCTTAGGAAACCCAGAATGCCAATTACTGATTTCCAAGGAGTCATGGTC
 CTACATTGTAGAAAAACCAAATCCTGAGAATGGAACATGTTACCCAGGGCATTTCGCTGACTATGAGGAACTGAGGG
 AGCAATTGAGTTCAGTATCTTCATTTGAGAGGTTTCAAATATTCCCCAAAGAAAGCTCATGGCCCAACACACCGTA
 ACCGGAGTGTGAGCATCATGCTCCCATATGGGGAAGCAGTTTTACAGAAATTTGCTATGGCTGACGGGGAAGA
 ATGGTTTTGATCCCAAACCTGAGCAAGTCCATGCAAAACAAAGAAAAAGAAAGTCTTGACTATGGGGTGTTCAT
 CACCCGCCAAACATAGGTGACCAAAAGGCCCTCTATCATACAGAAAATGCTTATGTCTGTAGTGTCTTCACATTAT
 AGCAGAAAATTCACCCAGAAATAGCCAAAAGACCCAAAGTAAGAGATCAAGAAGGAAGAATCAATTACTACTGGAC
 TCTGCTTGAACCCGGGGATACAATAATTTGAGGCAAATGGAATCTAATAGCGCCAAGATATGCTTTCGCACTGA
 GTAGAGGCTTTGGATCAGGAATCATCAACTCAAATGCACCAATGGATAAATGTGATGCCAAGTGCCAAACACCTCAG
 GGAGCTATAAACAGCAGTCTTCTTTCCAGAACGTACACCCAGTCACAATAGGAGAGTGTCCAAAGTATGTCAGGAG
 GCAAAAATTAAGGATGGTTACAGGACTAAGGAACATCCCATCCATTCAATCCAGAGGTTTGTTTGGAGCCATTGCCG
 GTTTCATTGAAGGGGGGTGGACTGGAATGGTAGATGGTTGGTATGGTTATCATCATCAGAATGAGCAAGGATCTGG
 CTATGCTGCAGATCAAAAAAGCACACAAAATGCCATTAATGGGATTACAAACAAGGTCAATTCTGTAATTGAGAAAAT
 GAACACTCAATTCACAGCAGTGGGCAAAGAGTTCAACAAATGGAAAGAAGGATGGAAAACCTGAATAAAAAAGTTG
 ATGATGGGTTTATAGACATTTGGACATATAATGCAGAACTGTTGGTTCTACTGGAAAATGAAAGGACTTTGGATTTC
 ATGACTCCAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAGTTAAAGAATAATGCTAAAGAAATAGGAAATGGGT
 GTTTTGAGTTCTATCACAAGTGAACGATGAATGCATGGAGAGTGAAGAATGGAACCTTATGACTATCCAAAATATT
 CCGAAGAATCAAAGTTAAACAGGGAGAAAATTGATGGAGTGAAATTGGAATCAATGGGAGTCTATCAGATTCTGGCG
 ATCTACTCAACAGTCGCCAGTTCTCTGGTTCTTTTGGTCTCCCTGGGGGCAATCAGCTTCTGGATGTGTTCCAATGG
 GTCTTTACAGTGTAGAATATGCATCTAAAGGCCTATTTCTTTAGTTTGAATTTACTGTTATTGCGGTGTGCATTTCTAT
 GTTTGGTGAGCGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTACGAGG
 TCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTAAGAAAGGACCGGGAATTCGATATCAA
 GCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATG
 ATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGG
 TTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAAT
 ATCGCGCGCGGTGTCATCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

SEQ. ID. NO. 69

TAAATTAGAAATTCGAGCTCCACCGCGGAAACCTCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAAG
ATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTG
CCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTT
CAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGAATGACGCACAATCCCACTATCCTTCGCAAGACCC
TTCTCTATATAAGGAAGTTCATTTTCATTTGGAGAGTATTAATACTTAATAGGTTTTGATAAAAAGCGAACGTGGG
AAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTTGTCTTTCTTGCCTGAG
CGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAGATCTCT
TTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCTTCTTGTCTGGTCTTGGGAAAAGAA
AGCTTGCTGGAGGCTGCTGTTACAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCT
CTACTTCTGCTTGACGAGGTATTGCTGCTGTACTTCTTCTTCTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAG
TATTTTCTTTGAAACAGAGTTTTCCCGTGTTTTGCAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTTCTGCCCAA
ATTTGTCGGGCCCCATGGAGAAAATAGTGCTTCTTCTTGCATAATGACGTCTTGTAAAAGTGATCAGATTTGCATTG
GTTACCATGCAAACAATTCAACAGAGCAGGTTGACACAATCATGGAAAAAGAACGTTACTGTTACACATGCCCAAGA
CATACTGGAAAAGACACACAACGGGAAGCTCTGCGATCTAGATGGAGTGAAGCCTCTAATTTTAAGAGATTGTAGT
GTAGCTGGATGGCTCCTCGGGAACCCAATGTGTGACGAATTCATCAATGTACCGGAATGGTCTTACATAGTGGAGA
AGGCCAATCCAACCAATGACCTCTGTTACCCAGGGCATTTCCGTGACTATGAGGAACCTGAGGGAGCAATTGAGTTC
AGTATCTTCATTTGAGAGTTGCAAAATATCCCCAAAGAAAGCTCATGGCCCAACCAACCCGTAACCGGAGTGTC
GCATCATGCTCCCATAAATGGGGAAAGCAGTTTTTACAGAAATTTGCTATGGCTGACGCGGGAAGAATGGTTTGTACC
CAAACCTGAGCAAGTCTATGCAAACAACAAAGAAAAAGAAGTCTTGACTATGGGGTGTTCATCACCCGCCCAA
CATAGGTGACCAAAAGGCCCTCTATCATACAGAAAATGCTTATGTCTCTGTAGTGTCTTCACATTATAGCAGAAAAT
TCACCCAGAAAATAGCCAAAAAGCCCAAAGTAAGAGATCAAGAAGGAAGAATCAATTACTACTGGACTCTGCTTGA
ACCCGGGGATACAATAATTTGAGGCAATGGAATCTAATAGCGCCAAGATGCTTTCCGCACTGAGTAGAGGC
TTTGATCAGGAATTATGAAAAGTGAATTTGGAATATGGTAAGTGAACCAACCAAGTGCAAACTCCAATGGGGCGA
TAAACTCTAGTATGCCATTCCACAACATACACCCTCTCACCATCGGGGAATGCCCCAAATATGTGAAATCAAAAGCA
TTAGTCTTGAACAGGGCTCAGAAATAGCCCTCAAAGAGAGAGCAGAAGAAAAAGAGAGGACTATTTGGAGCTA
TAGCAGGTTTTATAGAGGGAGGATGGCAGGGAATGGTAGATGGTTGGTATGGGTACCACCATAGCAATGAGCAGG
GGAGTGGGTACGCTGCAGACAAGAATCCACTCAAAAGGCAATAGGTGGAGTCAACCAATAGGTCAACTCAATCAT
TGACAAAATGAACACTCAGTTTGAGGCCGTTGGAAGGGAATTTAATAACTTAGAAAGGAGATAGAGAATTTAAACA
AGAAGATGGAAGACGGGTTTTCTAGATGTCTGGACTTATAATGCCGAACCTCTGGTTCTCATGGAAAATGAGAGA
TCTAGACTTTTCATGACTCAAATGTTAAGAACCTCTACGACAAGGTCCGACTACAGCTTAGGGATAATGCAAGGAG
CTGGGTAAACGGTTGTTTCGAGTTCTATCACAAATGTGATAATGAATGTATGGAAAGTATAAGAAACGGAACGTACAA
CTATCCGCAGTATTCAGAAGAAGCAAGATTAAGAAAGAGAGGAAATAAGTGGGGTAAAATTGGAATCAATAGGAAC
TACCAAACTACTGTCAAATTTATTCACAGTGCGGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATG
GATGTGCTCCAATGGATCGTTACAATGCAGAAATTTGCATTTAAAGGCCATATTTTCTTGTGATTTGATTTACTGTTAT
CGGTGTGCAATTTCTATGTTTGGTGAGCGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTG
AGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAAGAAAAAAGAAAAAAGAA
CCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAACATTTGGCAATAAAGTTTCTTAAGATTGAATCC
TGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATACGTTAAGCATGTAATAATTAACATGTAATGCAT
GACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATA
GCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCG
CG

Fig. 35

SEQ. ID. NO. 70

Đoạn DraIII-Plasto(-84+1)-H3 A/Brisbane/10/07-SacI được tổng hợp

CACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAAT
 CATCTTGAGAGAAAATGAAGACTATCATTGCTTTGAGCTACATTCTATGTCTGGTTTTCACTCAAAAAC
 TCCCGGAAATGACAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATA
 GTGAAAACAATCACGAATGACCAAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAGTTCCTCAACA
 GGTGAAATATGCGACAGTCCTCATCAGATCCTTGATGGAGAAAACGCACACTAATAGATGCTCTATTG
 GGAGACCCTCAGTGTGATGGCTTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGCAGCAAAGCCTA
 CAGCAACTGTTACCCTTATGATGTGCCGGATTATGCCTCCCTTAGGTCACTAGTTGCCTCATCCGGCA
 CACTGGAGTTTAAACAATGAAAGTTTCAATTGGACTGGAGTCACTCAAAACGGAACAAGCTCTGCTTGC
 ATAAGGAGATCTAATAACAGTTTCTTTAGTAGATTGAATTGGTTGACCCACTTAAAATTCAAATACCCAG
 CATTGAACGTGACTATGCCAAACAATGAAAAATTTGACAAATTGTACATTTGGGGGGTTCCACCACCCG
 GGTACGGACAATGACCAAATCTTCCTGTATGCTCAAGCATCAGGAAGAATCACAGTCTCTACCAAAG
 AAGCCAACAACTGTAATCCCGAATATCGGATCTAGACCCAGAGTAAGGAATATCCCCAGCAGAATAA
 GCATCTATTGGACAATAGTAAAACCGGGAGACATACTTTTGATTAACAGCACAGGGAATCTAATTGCTC
 CTAGGGGTTACTTCAAAATACGAAGTGGGAAAAGCTCAATAATGAGATCAGATGCACCCATTGGCAAA
 TGCAATTCTGAATGCATCACTCCAAACGGAAGCATTCCCAATGACAAACCATTCCAAAATGTAAACAGG
 ATCACATACGGGGCCTGTCCCAGATATGTTAAGCAAAACACTCTGAAATTGGCAACAGGGATGCGAAA
 TGTACCAGAGAAACAACTAGAGGCATATTTGGCGCAATCGCGGGTTTCATAGAAAATGGTTGGGAGG
 GAATGGTGGATGGTTGGTATGGTTTCAGGCATCAAATTCTGAGGGAATAGGACAAGCAGCAGATCTC
 AAAAGCACTCAAGCAGCAATCGATCAAATCAATGGGAAGCTGAATAGGTTGATCGGGAAAACCAACGA
 GAAATTCCATCAGATTGAAAAAGAGTTCTCAGAAGTCGAAGGGAGAATCCAGGACCTTGAGAAATATG
 TTGAGGACACCAAATAGATCTCTGGTCATACAACGCGGAGCTTCTTGTTGCCCTGGAGAACCAACAT
 ACAATTGATCTAACTGACTCAGAAATGAACAACTGTTTGAAAAAACAAAGAAGCAACTGAGGGAAAAT
 GCTGAGGATATGGGCAATGGTTGTTTCAAAATATACCACAAATGTGACAATGCCTGCATAGGATCAATC
 AGAAATGGAACCTTATGACCACGATGTATACAGAGATGAAGCATTAAACAACCGGTTCCAGATCAAGGG
 CGTTGAGCTGAAGTCAGGATACAAAGATTGGATACTATGGATTTCTTTGCCATATCATGTTTTTTGCT
 TTGTGTTGCTTTGTTGGGGTTCATCATGTGGGCCTGCCAAAAGGCAACATTAGGTGCAACATTTGCA
 TTTGAGAGCTC

Fig. 36

SEQ. ID. NO. 71

Sản phẩm thiết kế 736, từ PacI (ngược chiều với trình tự khởi đầu 35S) đến AseI (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa SpPDI-H3 A/Brisbane/10/2077 được gạch dưới.

TTAATTAAGAATTCGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAA
 GATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTC
 TGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACGT
 CTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAG
 ACCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATAATCTTAATAGTTTTGATAAAGCGAACG
 TGGGGAACCCGAACCAACCTTCTCTAACTCTCTCATCTCTCTTAAGCAAACCTTCTCTTGTCTTTCTTG
 CGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAA
 GATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGG
 AAAAGAAAGCTTGCTGGAGGCTGCTGTTGAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTG
 CAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTGCTGATTGGTTCTATAAG
 AAATCTAGTATTTTTCTTGAACAGAGTTTTCCCGTGGTTTTCGAACTTGAGAAAGATTGTTAAGCTTCTGTATATT
 CTGCCCCAAATTTGTGCGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTC
 CTTCTCAGATCTTCGCTCAAAAACCTCCCGGAAATGACAAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCGAG
 TACCAACCGGAACGATAGTGAAAAAATCACGAATGACCAATTGAAGTTACTAATGCTACTGAGCTGGTTCAGAG
 TTCCTCAACAGGTGAAATATGCGACAGTCTCATCAGATCCTTGATGGAGAAAACTGCACACTAATAGATGCTCTA
 TTGGGAGACCCTCAGTGTGATGGCTTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGCAGCAAAGCCTACAGC
 AACTGTTACCCCTTATGATGTGCCGGATTATGCCTCCCTTAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTTTA
 ACAATGAAAGTTTCAATTGGACTGGAGTCACTCAAAACGGAACAAGCTCTGCTTGATAAGGAGATCTAATAACAG
 TTTCTTTAGTAGATTGAATTGGTTGACCCACTTAAATTCAAATACCCAGCATTGAACGTGACTATGCCAAACAATG
 AAAAATTTGACAAATTGTACATTTGGGGGGTTTACCACCCGGGTACGGACAATGACCAATCTTCTGTATGCTCA
 AGCATCAGGAAGAATCACAGTCTCTACCAAAAAGAAGCCAACAACTGTAATCCCGAATATCGGATCTAGACCCAG
 AGTAAGGAATATCCCCAGCAGAATAAGCATCTATTGGACAATAGTAAACCCGGGAGACATACTTTTGATTAAACAGC
 ACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAAATACGAAGTGGGAAAAGCTCAATAATGAGATCAGATGCAC
 CCATTGGCAAATGCAATTCTGAATGCATCACTCCAAACGGAAGCATTCCCAATGACAAACCATTCCAAAATGTAAG
 CAGGATCACATACGGGGCTGTCCCAGATATGTTAAGCAAAACACTCTGAAATTGGCAACAGGGATGCGAAATGT
 ACCAGAGAAACAACTAGAGGCATATTTGGCGCAATCGCGGGTTTCATAGAAAATGGTTGGGAGGGAATGGTGG
 ATGGTTGGTATGGTTTCAGGCATCAAAATCTGAGGGAATAGGACAAGCAGCAGATCTCAAAAGCACTCAAGCAG
 CAATCGATCAATCAATGGGAAGCTGAATAGGTTGATCGGGAAAACCAACGAGAAATCCATCAGATTGAAAAAGA
 GTTCTCAGAAAGTGAAGGGGAGAATCCAGGACCTTGAGAAATATGTTGAGGACACCAAAATAGATCTCTGGTCATA
 CAACGCGGAGCTTCTTGTGGCCTGGAGAACCAACATACAATTGATCTAACTGACTCAGAAATGAACAAACTGTTT
 GAAAAACAAAGAAAGCAACTGAGGGAAAATGCTGAGGATATGGGCAATGTTGTTTCAAAATATACCACAAATGTG
 ACAATGCCTGCATAGGATCAATCAGAAATGGAACCTTATGACCACGATGTATACAGAGATGAAGCATTAAACAACCG
 GTTCCAGATCAAGGGCGTTGAGCTGAAGTCAGGATACAAAGATTGGATACTATGGATTTCTTTGCCATATCATGT
 TTTTGTCTTGTGTTGCTTTGTTGGGGTTTCATCATGTGGGCCTGCCAAAAAGGCAACATTAGGTGCAACATTTGCA
 TTTGAAGGCCTATTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTG
 CTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCTGCCCTCAGCAAGGACA
 CAAAAAGATTTTAAATTTTATTAATAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGAT
 CGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCATGATTATCATATAATTTCT
 GTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGT
 CCCGCAATTATACATTTAATACGCGATAGAAAACAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTG
 TCATCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Fig. 37

SEQ. ID. NO. 72

Sản phẩm thiết kế 737, từ PacI (ngược chiều với trình tự khởi đầu 35S) đến Ascl (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa HA thể khảm được gạch dưới.

TTAATTAAGAAATTCGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAAGA
TAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGC
CGACAGTGGTCCCAAGATGGACCCCCACCCACGAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACGTCTTCA
AAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCCTTC
CTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATACTTAATAGGTTTTGATAAAGCGAACGTGGGGAAA
CCCGAACCAAAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTTGTCTTTCTTGCCTGAGCGAT
CTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAGATCTCTTTGTG
GACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGGAAAAGAAAGCTTG
CTGGAGGCTGCTGTTGAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCT
GCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTT
GAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCAAATTTGTGCGG
CCCATGGCGAAAAACGTTGCGATTTTTCGGCTTATTGTTTTCTTCTTGTGTTGGTTCTTCTCAGATCTTCGCTCAA
AACTTCCCGGAAATGACAACAGCACGGCAACGCTGTGCCTTGGGCACCATGCAGTACCAAACGGAACGATAGTGA
AAACAATCACGAATGACCAAATGAAGTTACTAATGCTACTGAGCTGGTTGAGAGTTCTCAACAGGTGAAATATGCG
ACAGTCTCATCAGATCCTTGATGGAGAAAACTGCACACTAATAGATGCTCTATTGGGAGACCCCTCAGTGTGATGGC
TTCCAAAATAAGAAATGGGACCTTTTTGTTGAACGCGAGCAAAGCCTACAGCAACTGTTACCTTATGATGTGCCGGAT
TATGCCTCCCTTAGGTCACTAGTTGCCTCATCCGGCACACTGGAGTTTAAACAATGAAAGTTTCAATTGGACTGGAGT
CACTCAAAACGGAACAAGCTCTGCTTGCATAAGGAGATCTAATAACAGTTTTCTTAGTAGATTGAATTGGTTGACCCA
CTTAAATTCAAATACCCAGCATTGAACGTGACTATGCCAAACAATGAAAAATTTGACAAATTTGTACATTTGGGGGGT
TCACCACCCGGGTACGGACAATGACCAAATCTTCTGTATGCTCAAGCATCAGGAAGAATCACAGTCTCTACCAAAA
GAAGCCAACAACTGTAATCCCGAATATCGGATCTAGACCCAGAGTAAGGAATATCCCCAGCAGAATAAGCATCTAT
TGGACAATAGTAAAACCGGGAGACATACTTTTGATTAAACAGCACAGGGAATCTAATTGCTCCTAGGGGTTACTTCAAA
ATACGAAGTGGGAAAAGCTCAATAATGAGATCAGATGCACCCATTGGCAAATGCAATTCTGAATGCATCACTCCAAA
CGGAAGCATTCCCAATGACAAACCATTCAAAATGTAACAGGATCACATACGGGGCCTGTCCAGATATGTTAAGC
AAAACACTCTGAAATTTGGCAACAGGGATGCGAAATGTACCAGAGAAAACAACTAGAGGCATATTTGGCGCAATCGCG
GGTTTCATAGAAAATGGTTGGGAGGGAATGGTGGATGGTTGGTATGGTTTCAGGCATCAAAATCTGAGGGAATAG
GACAAGCAGCAGATCTCAAAAGCACTCAAGCAGCAATCGATCAAAATCAATGGGAAGCTGAATAGGTTGATCGGGAA
AACCAACGAGAAATTCATCAGATTGAAAAAGAGTTCTCAGAAGTCAAGGGGAGAATCCAGGACCTTGAGAAATATG
TTGAGGACACCAAAATAGATCTCTGGTCATACAACGCGGAGCTTCTTGTGCCCCTGGAGAACCAACATACAATTGAT
CTAACTGACTCAGAAATGAACAACTGTTTGA AAAAACAAGCAACTGAGGGAAAATGCTGAGGATATGGGCAA
TGGTTGTTTCAAAATATACCACAAATGTGACAATGCCTGCATAGGATCAATCAGAAATGGAACCTTATGACCACGATGT
ATACAGAGATGAAGCATTAACAACCGGTTCCAGATCAAGGGCGTTGAGCTGAAGTCAATAGGAACCTTACCAAAATAC
TGTCATTTTATTCAACAGTGGCGAGTTCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCA
ATGGATCGTTACAATGCAGAATTTGCATTTAAAGGCCATTTTCTTTAGTTTGAATTTACTGTTATTCGGGTGTCATTT
CTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAG
CAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTA AAAAAAAAAAAAAAAAAAGACCGGAATTCGAT
ATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTG
CGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGA
TGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAATATAGCGCGCAAACTAGGAT
AAATTATCGCGCGCGGTGTCTATGTTACTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Fig. 38

SEQ. ID. NO. 73

DraIII-Plasto(-84+1)-HA B/Florida/4/06-SacI được tổng hợp

CACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAA
TCATCTTGAGAGAAAATGAAGGCAATAATTGTACTACTCATGGTAGTAACATCCAATGCAGATCGAAT
CTGCACTGGAATAACATCTTCAAACCTCACCTCATGTGGTCAAAACAGCCACTCAAGGGGAGGTCAAT
GTGACTGGTGTGATACCACTAACAACAACACCAACAAAATCTTATTTTGCAAATCTCAAAGGAACAAG
GACCAGAGGGAACTATGCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACC
AATGTGTGTGGGGACCACACCTTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTACATCC
GGGTGCTTTCTATAATGCACGACAGAACAAAATCAGGCAACTACCCAATCTTCTCAGAGGATATGA
AAATATCAGGCTATCAACCCAAAACGTATCGATGCGGAAAAGGCACCAGGAGGACCCTACAGACTT
GGAACCTCAGGATCTTGCCCTAACGCTACCAGTAAGAGCGGATTTTTCGCAACAATGGCTTGGGCTG
TCCCAAAGGACAACAACAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACATTTGTACAGAA
GGGGAAGACCAAATCACTGTTTGGGGTTCCATTCAGATAACAAAACCCAAATGAAGAACCTCTATG
GAGACTCAAATCCTCAAAAGTTCACCTCATCTGCTAATGGAGTAACCACACACTATGTTTCTCAGATT
GGCAGCTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAAGCGGCAGGATTGTTGTTGATTACA
TGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTACCAAAGAGGTGTTTTGTTGCCTCAAAAGGTG
TGGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCTTTAATTGGTGAAGCAGATTGCC
TTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATA
GGAAATTGCCCAATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAAATATAGACCTCCTG
CAAAACTATTAAAGGAAAGGGGTTTTCTTCGGAGCTATTGCTGGTTTCCTAGAAGGAGGATGGGAAGG
AATGATTGCAGGCTGGCACGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGCGGACCT
TAAGAGTACGCAAGAAGCTATAAACAAGATAACAAAAAATCTCAATTCTTTGAGTGAGCTAGAAGTAAA
GAATCTTCAAAGACTAAGTGGTGCCATGGATGAACTCCACAACGAAATACTCGAGCTGGATGAGAAA
GTGGATGATCTCAGAGCTGACACTATAAGCTCGCAAATAGAACTTGCAGTCTTGCTTTCCAACGAAG
GAATAATAAACAGTGAAGATGAGCATCTATTGGCACTTGAGAGAAAATAAAGAAAATGCTGGGTCCC
TCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAACACAAGTGCAACCAGACCTGCTTAGACA
GGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCCACTTTTGATTCACTGAACATTACT
GCTGCATCTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTTCT
AGTTTGGCTGTAAACATTGATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGC
TCCATCTGTCTATAAGAGCTC

SEQ. ID. NO. 74

TTAATTAAGAATTCCGAGCTCCACCGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAA
GATAGTGGAAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTC
TGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGT
CTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCAACAATCCCCTATACCTTCGCAAG
ACCTTTCCTCTATATAAGGAAGTTCATTTCACTTGGAGAGGTATTAATACTTAATAGTTTTGATAAAAGCGAACG
TGGGGAAACCCGAACCAACCTTCTTCTAAACTCTCTCATCTCTCTTAAAGCAAACTCTCTCTTGTCTTTCTTG
CGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAA
GATCTCTTTGTGGACACGTAGTGCGGCGCCATTAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGG
AAAAAGAAAGCTTGCTGGAGGCTGCTGTTACGCCCCATACATTACTTGTTACGATTCTGCTGACTTTGCGCGGGTG
CAATATCTCTACTTCTGCTTGACGAGGATTTGTTGCCTGTACTTCTTCTTCTTCTTCTTCTTCTGATTGGTTCTATAAG
AAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGTTTTCGAACTGGAGAAAGATTGTAAGCTTCTGTATATT
TGCCCCAAATTTGTGCGGGCCCATGGCGAAAAACGTTGCGATTITTCGGCTTATTGTTTTCTTCTTCTTGTGGTTTC
CTTCTCAGATCTTCGCTGATCGAATCTGCACTGGAATAACATCTTCAAACCTCACCTCATGTGGTCAAAACAGCCAC
TCAAGGGGAGGTCAATGTGACTGGTGTGATACCACTAACAACAACACCAACAAAATCTTATTTGCAATCTCAAA
GGAACAAGGACCAGAGGGGAAACTATGCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACC
AATGTGTGTGGGGACCACACCTTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTACATCCGGGTGCTT
TCCTATAATGCACGACAGACGAAAAATCAGGCAACTACCCAATCTTCTCAGAGGATATGAAATATCAGGCTATCA
ACCCAAAAATGCTATCGATGCAAGGAAAAAGGACCAGGAGGACCCCTACAGACTTGAACCTCAGGATCTGCCCTAA
CGCTACCAGTAAGAGCGGATTTTTCGCAACAATGGCTTGGGCTGTCCCAAAGGACAACAACAAAATGCAACGAA
CCCCTAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAAATCACTGTTTGGGGGTTCCATTGAGAT
AACAAAACCCCAAATGAAGAACCTCTATGGAGACTCAAATCCTCAAAAGTTACCTCATCTGCTAATGGAGTAACCA
CACACTATGTTTCTCAGATTGGCAGCTTCCCAGATCAACACGAAGACGGAGGACTACCACAAGCGCGCAGGATTG
TGTGTTGATTACATGATGCAAAAACCTGGGAAACAGGAACAATTGTCTACCAAGAGGTTGTTTTGTGCTCAAAA
GGTGTGGTGCAGGATGGCAGGAGCAAAAGTAATAAAAGGGTCTTGCCTTTAATTGGTGAAGCAGATTGCTCTTCA
TGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAAATTGCC
AATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAAATATAGACCTCCTGCAAACTATTAAGGAAAGG
GGTTTCTTCGGAGCTATTGCTGGTTTCTAGAAGGAGGATGGGAAGGAATGATTGCAGGCTGGCACGGATACAC
ATCTCACGGAGCACATGGAGTGGCAGTGGCGCGGACCTTAAGAGTACGCAAGAGCTATAAACAAGATAACAA
AAAACTCAATTTCTTTGAGTGAGCTAGAAGTAAGAAATCTTCAAAGACTAAGTGGTCCATGGATGAACCTCCACA
CGAAATACTCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCGCAAATAGAAGTTCGAGT
CTTGCTTTCCAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAATG
CTGGGTCCCTCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAACACAAGTGAACCCAGACCTGCTTAGAC
AGGATAGCTGCTGGCACCTTTAATGCAGGAGAAATTTCTCTCCCCACTTTTGATTCACTGAACATTACTGCTGCAT
CTTTAAATGATGATGGATTGGATAACCATACTATACTGCTCTATTACTCAACTGCTGCTTCTAGTTTGGCTGTACA
TTGATGCTAGCTATTTTTATTGTTTATATGGTCTCCAGAGACAACGTTTCATGCTCCATCTCTATAAAGGCCAT
TTTCTTTAGTTTGAATTTACTGTTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTG
TTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTGTCCTTCAGCAAGGACACAAAAAGATTTTA
ATTTTATTAATAAAAAAAAAAAAAAAAAAAGACCGGAATTGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTT
GGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCATGATTATCATATAATTTCTGTTGAATTACGTT
AAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGAATTATA
CATTTAATACGCGATAGAAAAACAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTA
CTAGATTCTAGAGTCTCAAGCTTCGGCGCGCC

Fig. 40

SEQ. ID. NO. 75

Sản phẩm thiết kế 745, từ Pacl (ngược chiều với trình tự khởi đầu 35S) đến Ascl (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa HA thể khảm được gạch dưới.

TTAATTAAGAATTGAGCTCCACGCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGAGAA
 GATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTC
 TGCCGACAGTGGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACGT
 CTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGA
 CCCTTCCTCTATATAAGGAAGTTCATTTTCAATTTGGAGAGGTATTAATAATCTTAATAGGTTTTGATAAAGCGAACGT
 GGGGAAACCCGAACCAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAGCAAACCTCTCTCTTGTCTTTCTTGC
 GTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAG
 ATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGGA
 AAAGAAAAGCTTGCTGGAGGCTGCTGTTAGCCCCATACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCA
 ATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTCTTCTTCTTCTGCTGATTGGTTCTATAAGAA
 ATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCT
 GCCCAAATTTGTGCGGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCTT
 CTCAGATCTTCGCTGATCGAATCTGCACTGGAATAACATCTTCAAACCTCACCTCATGTGGTCAAACAGCCACTCA
 AGGGGAGGTCAATGTGACTGGTGTGATACCACTAACCAACACCAAAATCTTATTTTGCAAAATCTCAAAGGA
 ACAAGGACCAGAGGGAACTATGCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCTTTGGGCAGACCAATG
 TGTGTGGGGACCACACCTTCGGCGAAGGCTTCAATACTCCACGAAGTCAAACCTGTTACATCCGGGTGCTTTCTC
 ATAATGCACGACAGAACAAAAATCAGGCAACTACCAATCTTCTCAGAGGATATGAAAATATCAGGCTATCAACCC
 AAAACGTCATCGATGCGGAAAAAGGCACCAGGAGGACCCTACAGACTTGGAACTCAGGATCTTGCCCTAACGCTA
 CCAGTAAGAGCGGATTTTTCGCAACAATGGCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCAC
 TAACAGTAGAAGTACCATACATTTGTACAGAAGGGGAAGACCAAAATCACTGTTTGGGGGTTCCATTAGATAACAA
 AACCCAAATGAAGAACCTCTATGGAGACTCAAATCCTCAAAAGTTTCACTCATCTGCTAATGGAGTAACCAACAC
 TATGTTTCTCAGATTGGCAGCTTCCCAGATCAAACAGAAGACGGAGGACTACCACAAAGCGGCAGGATTGTTGTT
 GATTACATGATGCAAAAACCTGGGAAAACAGGAACAATTGTCTACCAAAAGAGGTGTTTTGTTGCCTCAAAAGGTGT
 GGTGCGCGAGTGGCAGGAGCAAAGTAATAAAAGGGTCCTTGCCCTTAATTGGTGAAGCAGATTGCCCTTCATGAAA
 AATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAACATGCAAAAGCCATAGGAAATTGCCCAATATG
 GGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAATATAGACCTCCTGCAAAACTATTAAAGGAAAGGGGTTTC
 TTCGGAGCTATTGCTGGTTTCTAGAAGGAGGATGGGAAGGAATGATTGCAGGCTGGCACGGATACACATCTCAC
 GGAGCACATGGAGTGGCAGTGGCGGCGGACCTTAAGAGTACGCAAGAAGCTATAAACAAGATAACAAAAAATCTC
 AATTCTTTGAGTGAGCTAGAAGTAAAGAATCTTCAAAGACTAAGTGGTGCCATGGATGAACCTCCACAACGAAATAC
 TCGAGCTGGATGAGAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCGCAAATAGAACTTGCAGTCTTGCTTTC
 CAACGAAGGAATAATAAACAGTGAAGATGAGCATCTATTGGCACTTGAGAGAAAACCTAAAGAAAATGCTGGGTCCC
 TCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAACACAAGTGCAACCAGACCTGCTTAGACAGGATAGCT
 GCTGGCACCTTTAATGCAGGAGAATTTTCTCTCCCACTTTTGATTCACTGAACATTACTGCTGCATCTTTAAATGA
 TGATGGATTGGATAACTACCAAACTGTCAATTTATTCAACAGTGGCGAGTTCCTAGCACTGGCAATCATGATG
 GCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAGAATTTGCATTTAAAGGCCATTTTTCTTTAG
 TTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATG
 TAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTAA
 AAAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAG
 TTTCTTAAGATTGAATCCTGTTGCCGGTCTTGGCATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAAT
 AATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCAGCAATTATACATTTAATACGC
 GATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATTCTAGA
 GTCTCAAGCTTCGGCGCGCGC

Fig. 41

SEQ. ID. NO. 76

Trình tự mã hóa Msj1.

ATGTTTGGGCGCGGACCAACAAGGAAGAGTGATAACACCAAATATTACGATATTCTTGGTGTTCAAA
AAGTGCTAGTGAAGATGAAATCAAGAAAGCCTATAGAAAGGCAGCGATGAAGAACCATCCAGATAAG
GGTGGGGATCCTGAGAAGTTCAAGGAGTTGGGCCAAGCATATGAAGTGTTGAGCGATCCTGAAAAG
AAAGAACTGTATGATCAATATGGTGAAGATGCCCTTAAAGAAGGAATGGGGGGAGGCGCAGGAAGC
TCATTTTCATAATCCGTTTGATATTTTCGAATCATTTTTTGGTGCAGGCTTTGGTGGTGGTGGTCCTTCA
CGCGCAAGAAGACAGAAGCAAGGAGAAGATGTGGTGCATTCTATAAAGGTTTCCTTGGAGGATGTGT
ATAACGGCACTACAAAGAAGCTATCACTTTCTAGGAATGCACTGTGCTCAAAATGTAAAGGGAAAGGT
TCAAAAAGTGGAAGCTGCTGGAAGGTGTTTTGGATGCCAGGGCACAGGTATGAAGATTACCAGAAGG
CAAATTGGACTGGGCATGATTCAACAAATGCAACACGTCTGTCTGACTGCAAAGGAACAGGCGAGG
TCATTAGTGAGAGAGATAGATGCCCTCAATGCAAGGGAAACAAGATTACTCAAGAAAAGAAGGTGCT
GGAGGTGCATGTGGAAAAGGGGATGCAGCAGGGTCACAAGATTGTATTGCAAGGACAAGCTGATGA
AGCTCCTGATACAATCACAGGAGACATAGTTTTTGTCTTGCAAGTAAAGGGACATCCGAAGTTTCGGA
GGGAGCGTGATGACCTCCACATTGAACACAATTTGAGCTTAACTGAGGCTCTCTGTGGCTTCCAGTT
TAATGTCACACATCTTGATGGAAGGCAACTATTGGTCAAATCGAACCCCGGCGAAGTCATCAAGCCA
GGTCAACATAAAGCTATAAATGATGAGGGAATGCCACAACATGGTAGGCCGTTTCATGAAGGGACGCC
TATACATCAAGTTTAGTGTTGATTTCCCGGATTCGGGTTTTCTTTCCCCAAGCCAAAGCCTGGAATTA
GAAAAGATATTACCTCAAAGACAAGCAAGAACTTGTCCTCAAGGAGGAGGTAGATGATTGTGAGGAGA
CCACCCTGCATGATGTCAATATTGCAGAGGAGATGAGTCGAAAGAAGCAACAATACCGTGAGGCATA
TGATGACGATGATGATGAAGATGATGAGCACTCGCAGCCTCGGGTGCAATGCGCTCAACAGTAG

Fig. 42

39/49

SEQ. ID. NO. 77

Sản phẩm thiết kế số R850, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với promotor) đến EcoRI (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa HSP40 được gạch dưới.

AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGCTGGTCTGTACATTCATCTTGCCGCTTTGCATT
CACTTGGCCACAAAGAGTAGAGAGAAGGAAGAGAAGAGCCAGACTTCAAGAAGCGACCTTGCAAGTGCACCTCG
AGGGTCAGAACTGTATATCATATCTATGTGAGAGAAAGGGGAACATTTGAGATGGAGTCCATTTACTTGAGGTAT
ACTTATTATTTGATCAATAAATTTGTACTTCTTATTTAGATCAATAAATTTGTCATTAAGCTATAATCCAAAATAA
ATTACGATCAAAATATGCAAATGTTAGCCAGTACTTGTGTTAAACTTGATGGCATCTCTTGTTTCTTTGGCAATCAC
ATGCCTAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACTCTGTTTGGATAAAC
AGCTTAATTAAGCGCTTATAGAATATCATATGATTGTGTTTGGTCAGACTTCAGAGCATCTCTTGTTTCTCTGGCA
ATCATATGCCTAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACCCTGTTTGGG
TAAACAGCTTAATTAAGTGCTTATAGAATAAGCGCTTATCATATAAGTGCTTTTGACAGTTATTTCTATGAAAGTA
GAAGAAATAGTCATATTGTTTTAATAAGCTATCCTGGAGAGCTTGTTGGAATAACCAGAAAAGAACTTATGGAC
ACGTCATGAGCTGTTTACATAAGATCTCCCTAACAGTCTCAAAAGTGTTTATGCCAGTAGATAAATTCAAATAAGTC
AATCTAAACAGACCCCTAAATCCATTATGGTACCTATCATTTTAGCTTATTCCATCTTTATTAAGAATGTCATGAGATA
ACATAATGATAACACATTATTTTGACACAAATGGGCAGATCTAGCAATTTAACTCTGGAGTCCCTCAAGACTGCTG
TTCTTACGAAGTTCACGTCCCTGAATCATGTTCCCTGTATGGAAGCCTGAAAGACCTCAAATCTAAAAGGTGGCGA
TAAATTGAAGGTTTACAAAATATACCCTGCGGGCTTGACACAGAGGCAAGCTCTTTATACCTCCAGTTCAACGG
GGATGTTGATTTCAGAAGTCATTTGGAGAGCAATCCTTGTTGCCAAGTTTGAAGTAATTTTTGTGTAGCATATGTTG
AGCTACCTACAATTTACATGATCACCTAGCATTAGCTCTTTCACTTAACTGAGAGAATGAAGTTTTAGGAATGAGTA
TGACCATGGAGTCGGCATGGCTTTGTAATGCCTACCCTACTTTGGCCAACCTCATCGGGGATTTACATTGAGAAAA
TATACATGACTTCAACCATACTTAAACCCCTTTTTGTAAGATAACTGAATGTTCATATTTAATGTTGGGTTGTAGTG
TTTTTACTTGATTATATCCAGACAGTTACAAGTTGGACAACAAGATTGTGGGTCTGTACTGTTATTTATTTATTTTT
TTTTAGCAGAAACACCTTATCTTTGTTTCTGTTGAATGTAGAATGAAAATAAAGAAAGAAAATATAACATCATCG
GCCGCGCTTGCTAATTTGCGGCAGTTAGGATCCTCTCCGGTCACCGGAAAGTTTCAGTAGAAGAAACAAAACAC
CGTGACTAAAATGATACTATTATTTTATTGTTGTTTTCTTTTTCTACCGGAACTTTTAGAACGGATCCCAAC
TCGTTCCGGGGCCGCTACAACCTGAAACAAAAGAAGATATTTCTCTCTTTCAGAAATGTAAGTTTTCTTTACAG
ATACCCATTACCATTTGATTGAGATGTGGTGACTAGAGATAAAGCATACTAATTTGACTCTTGGAACCCATAAA
GTTTTATGTTATCCGTGTTCTGGACCAATCCACTTGGGGGCATAACCTGTGTCTATGTGTGGTTTGGTTCCATTCT
GATTTATGCGGCGACTTGTAATTTAAATCTAGGAGGGGCAGACATTGAACAATCCCAATATTTAATAACTTATG
CAAGATTTTTTTTATTAATGAGATGATGTGTTTGACTGAGATTGAGTCATACATTTCACTAAGAAATGGTTCCAA
GTACCAAACTATGATGACCCAGTTGCAAAACATGACGTTGCGGAGTGGTCACTTTGATAGTTCAATTTCACTTTGGC
TTCTTATTCCTTTTATAATTCTAATCTTCTGTGTAACTATTTTCATGTATTATTTTCTTTAAAATTTACATGTCATT
TATTTGCTCACTAATCAATTTGTCATATAACAATGATAAGTGATATTTGACTCACAAAATTTACATCAAATTTT
GACATCGTTTTATTATGTTTATTGGATGATTAACAAATATAACAACTTTGCAACTAATTAACCCCAACTGAATATAA
TTAACTATAACTGTGAAAGTAGTTAACTCATTTTTATTTTATAGATCAAATAAGAGAAATAACGGTATATTAATCC
CTCCAAAAAAGGAGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGG
ATAACATCCAAATCCAACTACACAACTCCTGATGATTAACCCACTTTAAGCCACGCATCTGTGGCAGATCT
ACATTATCTAAATCACACATTCTCCACACATCTGAGCCACACAAAACCAATCCACATCTTTATACCCATTCTAT
AAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTA
ATCATCTTGAGAGAAAATGTTTGGGCGCGGACCAACAAGGAAGAGTGATAACACCAAAATATTACGATATTCTTGG
TGTTTCAAAAAGTGCTAGTGAAGATGAAATCAAGAAAGCCTATAGAAAGGCAGCGATGAAGAACCATCCAGATAA
GGGTGGGGATCCTGAGAAGTTCAAGGAGTTGGCCAAGCATATGAAGTGTTGAGCGATCCTGAAAAGAAAGAAC
TGTATGATCAATATGGTGAAGATGCCCTTAAAGAAGGAATGGGGGGAGGCGCAGGAAGCTCATTTCAATATCCGT
TTGATATTTTGAATCATTTTTTGGTGCAGGCTTTGGTGGTGGTGGTCTTCACGCGCAAGAAGACAGAAGCAAG
GAGAAGATGTGGTGCATTCTATAAAGGTTTCTTGGAGGATGTGTATAACGGCACTACAAAGAAGCTATCACTTTC
TAGGAATGCACTGTGCTCAAAATGTAAAGGGAAAGGTTCAAAAAGTGGAAGTCTGGAAGGTGTTTTGGATGCCA
GGGCACAGGTATGAAGATTACCAGAAGGCAATTTGAGTGGGCATGATTCAACAAATGCAACACGTCTGTCTGTA
CTGCAAAGGAACAGGCGAGGTGATTAGTGAGAGAGATAGATGCCCTCAATGCAAGGGAAACAAGATTACTCAAG
AAAAGAAGGTGCTGGAGGTGCATGTGGAAGGGGATGCAGCAGGGTCACAAGATTGATTTCGAAGGACAAGCT
GATGAAGCTCCTGATACAATCACAGGAGACATAGTTTTTGTCTTGCAAGTAAAGGGACATCCGAAGTTTCGGAGG
GAGCGTGATGACCTCCACATTGAACACAATTTGAGCTTAACTGAGGCTCTCTGTGGCTTCCAGTTAATGTCACA
CATCTTGATGGAAGGCAACTATTGGTCAAATCGAACCCCGGCGAAGTCATCAAGCCAGGTCAACATAAAGCTATA
AATGATGAGGGAATGCCACAACATGGTAGGCCGTTTATGAAGGGACGCCTATACATCAAGTTTATGTTGATTTT
CCGGATTCCGGTTTTCTTTCCCAAGCCAAAGCCTGGAATTAGAAAAGATATTACCTCAAAAGACAAGCAAGAAC
TTGTCCCAAAAGGAGGTAGATGATTGTGAGGAGACCACCTGCATGATGTCAATATTGACAGAGGAGATGAGTCG
AAAGAAGCAACAATACCGTGAGGCATATGATGACGATGATGAAGATGATGAGCACTCGCAGCCTCGGGTGC
AATGCGCTCAACAGTAGGAGCTCAGCTCGAATTTCCCGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATT
GAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGT
AATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGAATTATACATTTAATACGCGATAGAAAA
AAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGGGTGTCATCTATGTTACTAGATCGAATTC

Fig. 43

40/49

SEQ. ID. NO. 78

Sản phẩm thiết kế số R860, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu đến EcoRI (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa HSP70 được gạch dưới.

AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGCTGGTCTGTACATTCATCTTGCCGCCTTTGCATTCACTTGGCCAC
AAAGAGTAGAGAGAAGGAAGAGAAGAGCCAGACTTCAAGAAGCGACCTTGAAGTGCCTCGAGGGTCAGAACTGTATATCATA
TCTATGTGAGAGAAAAGGGGAACATTTGAGATGGAGTCCATTTGACTGAGGTATACCTTATTATTTTGATCAATAAATTTGTACTTCTT
ATTTAGATCAATAAATTTGTCATTAAGCTATAATCCAAAATAAATTACGATCAAATATGCAAATGTTAGCCAGTACTTGTGTTAACTTG
ATGGCATCTCTTGGTTCTTTGGCAATCACATGCCTAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGAT
GACTCTGTTTGGATAAACAGCTTAATTAAGCGCTTATAGAATATCATATGATTGTGTTTGGTCAGACTTCAGAGCATCTCTTGGTTTCT
CTGGCAATCATATGCCTAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACCTGTTTGGGTAAACA
GCTTAATTAAGTGCTTATAGAATAAGCGCTTATCATATAAGTGCTTTGTACAGTTATTTCTATGAAAGTAGAAGAAATAGTCATATTG
TTTTAATATAAGCTATCCTGGAGAGCTTGTGGAAATAACCGAGAAAAGAACTTATGGACACGTCATGAGCTGTTTACATAAGATCTCCC
TAACAGTCTCAAAAGTGTTTATGCCAGTAGATAAATCAAATAAGTCAATCTAAACAGACCCTAAATCCATTATGGTACCTATCATTTT
AGCTTATTCCATCTTTTAAAGAATGTCATGAGATAACATAGATGATTGTTTGGTCAGACTTCAGAGTCAGATGACCTGTTTGGGTAAACA
TCTGGAGTCTTCAAGACTGCTGTTTACGAAGTTCAGTCCCTGAATCATGTTCTGTATGGAAGCCTGAAAGACCTCAAATTCTA
AAAGGTGGCGATAAATTGAAGGTTTACAAAATATACCTGCGGGCTTGACACAGAGGCAAGCTCTTTATACCTCCAGTTCAACGGG
GATGTTGATTTCAGAAGTCACTTGGAGAGCAATCCTGTGCCAAGTTTGAAGTAATTTTGTGTAGCATATGTTGAGCTACCTACAAT
TTACATGATCACGTACATTAAGCTCTTTCACCTTAATGACGAGAATGAAGTTTATGGAATGAGTATGACCATGGGACGTCGTCGCTTT
GTAATGCCTACCCTACTTTGGCCAACTCATCGGGGATTTACATTGAGAAAATATACATGACTTCAACCATACTTAAACCCCTTTTGTGA
AGATAACTGAATGTTTCAATTTAATGTTGGGTTGTAGTGTGTTTACTTGATTATATCCAGACAGTTACAAGTTGGACAACAAGATTGTG
GGTCTGACTGTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
AATATAACATCATCGGCCGCGCTTGTCTAATTTTCCGGGAGTTAGGATCCTCTCCGGTCACCGGAAAGTTTCAGTAGAAGAAACAAAA
CACCGTGACTAAAATGATACTATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT
GGGCCGCTACAACGAAACAAAAGAAGATATTTCTCTCTTTCAGAAATGAAGTTTCTTTACAGATACCCATTCCACCTTTGATT
CAGATGTGGTGACTAGAGATAAAGCATACTAATTTGACCTTGGAAACCCATAAAGTTTATGTTATCCGTTCTGGACCAATCCACT
TGGGGGCATAACCTGTGTCTATGTGTGGTTTGGTTTCCATTCTGATTTATGCGGCGACTTGAATTTAAAATTCAGAGGGGCGAGAC
ATTGAACAATCCCAATATTTTAACTTATGCAAGATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT
CACTAAGAAATGGTCCAAGTACCAAACTATCATGACCCAGTTGCAAAATGACGTTCCGGGAGTGGTCACTTTGATAGTTCAATTTT
ATCTGGCTCTTATTCTTTTATAATTCTAATCTCTTGTGTAACATTTTATGATTATTTTCTTTTAAAATTTACATGCTATTTATTT
TTGCCTCACTAATCAATTTGCATATAACAATGATAAGTGATTTTGAATCACAAAATTTACATCAAATTTGACATCGTTTATTTATG
TTCAATTTGATGATTAAACAAATATAACAACTTTGCAACTAATTAACCACCAACTGAATATAATTAACATAACTGTGAAAGTAGTTAAC
TCATTTTATATTTTATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAAGAGGATATTTACTAAAAATCTA
AGCCACGTAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCCAACCAATCAACAATCTGATGAGAGGATACCCACTT
TAAGCCACGCATCTGTGGCACATCTACATTTCTAATCACACATTCTCCACACATCTGAGCCACACAAAAACCAATCCACATCTT
TATCACCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAGAGACTAATTAATTA
TTAATCATCTTGAGAGAAAATGTCCGGTAAAGGAGAAGGACGAGCTATCGGTATCGATCTTGGTACCCTTACTCTTGGTCCGGAGT
ATGGTCAACGACGACCTGTTGAGATCATTGCTAATGATCAGGAAACAGAACCCGCTCTTACGTTGCTTTTACCGGACTCCGAGA
GGTTGATCGGTGACGCGAGCTAAGAATCAGGTCCGCTGAAACCCGCTTAAACCCGCTTTTTCGACGCTAAGAGGTTGATCGGTCTGCTG
TTCTCTGACAGCTCTGTTTCAAGTGTGACATGAAATTTGGGCCATTCAAGATTCAAGCCGGACCTGCCGATAGCCAATGATCTACGTC
GAATACAAGGGTGAAGAGAAAGAGTTCCGAGCTGAGGAGATTTCTCCATGGTCTTATTAAGATGCGTGAGATTGCTGAGGCTTA
CCTTGGTGTCAGAGCTGTTTGCCTACGGTGCTGCTGTCCAGGGAGCTATTCTACGCGGTGAAGGAAACGAGAAGGTTCAAGATCTTCT
CATCGCTGGTTTGAACGTTATGCGAATCATCAACGAGCCTACAGCCGCGGCTATTGCCTACGGTCTTGACAAAAAGGCTACCAAGCG
TTGGAGAGAAAGATGTTCTTATCTTCTGATCTTGGTGGTGGCACCTTTGATGCTCTCTTCTTACCATTGAAGAGGGTATCTTTGAGGT
GAAGGCAACTGCTGGTGACACCCATCTTGGTGGGGAAGATTTTGACAACAGAATGGTTAACCCTTTGTCCAAGAGTTCAAGAGGA
AGAGTAAGAAGGATATCACCGGTAACCCAAGAGCTCTTAGGAGGTTGAGAACTTCTGTGAGAGAGCGAAGAGGACTCTTTCTTCC
ACTGCTCAGACCACCATCGAGATTGACTCTCTATACGAGGGTATCGACTTCTACTCCACCATCACCCGCTGCTAGATTTGAGGAGCTC
AACATGGATCTCTTCAAGGAAGTGTATGGAGCCAGTTGAGAAAGTGTCTTCTGATGCTAAGATGGACAAGAGCACTGTTTATGATGTT
GTCTTGTGTTGGTGGTTTACCCGTATCCCTAAGGTTGAGCAATTTGCTCCAGGACTTCTTCAACGGCAAGAGCTTTGCAAGTCTATT
AACCTGATGAGGCTGTTTGCCTACGGTGCTGCTGTCCAGGGAGCTATTCTACGCGGTGAAGGAAACGAGAAGGTTCAAGATCTTCT
ATTGCTCGATGTCACTCTCTCTCCCTTGGTTTGGAACTGCCGGTGGTGTGATGACCACTTTGATCCCAAGGAACACAACCATCCC
AACCAAGAAAGAAACAGTCTTCTCCACCTACTCAGACAACCAACCCGGTGTGTTGATCCAGGTGTACGAAGGAGAGAGAGCCAGAA
CCAAGGACAACAACCTTCTTGGTAAATTTGAGCTCTCCGGAATTTCTCCAGCTCCTCGTGGTGTCCCCAGATCACAGTCTGCTTTG
ACATTGATGCCAATGGTATCCTCAATGTCTGCTGAGGACAAGCAACCCGACAGAGAAGAAAGATCAACCATCAACATGACAAG
GGTCTGCTCTCCAAGGATGAGATTGAGAAGATGGTTCAAGAGGCTGAGAAGTACAAGTCCGAAGACGAGGAGGACACAAGAAGAGG
TTGAAGCCAAGAACGCTCTCGAGAACTACGCTTACAACATGAGGAACACCATCCAAGACGAGAAGATTGGTGAGAAGCTCCCGGCT
GCAGACAAGAAGAAGATCGAGGATTCTATTGAGCAGGCGATTCAATGGCTCGAGGGTAACCAAGTTGGCTGAGGCTGATGAGTTCC
AAGACAAGATGAAGGAATTTGGAGAGCATCTGCAACCCAATCTTGCCAAGATGTACCAAGGAGCTGGTGGTGAAGCGGTGGTCC
AGGTGCCTCTGGTATGGACGATGATGCTCCCCCTGCTTCAAGGCGGTGCTGGACCTAAGATCGAGGAGGTCGACTAAGAGCTCAGC
TGAATTTCCCGATCGTTCAAACATTTGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATT
TCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAAT
TATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAACTAGGATAAATATCGCGCGCGGTGTCATCTATGTTACTAGATCG
AATTC

Fig. 44a

SEQ. ID. NO. 79

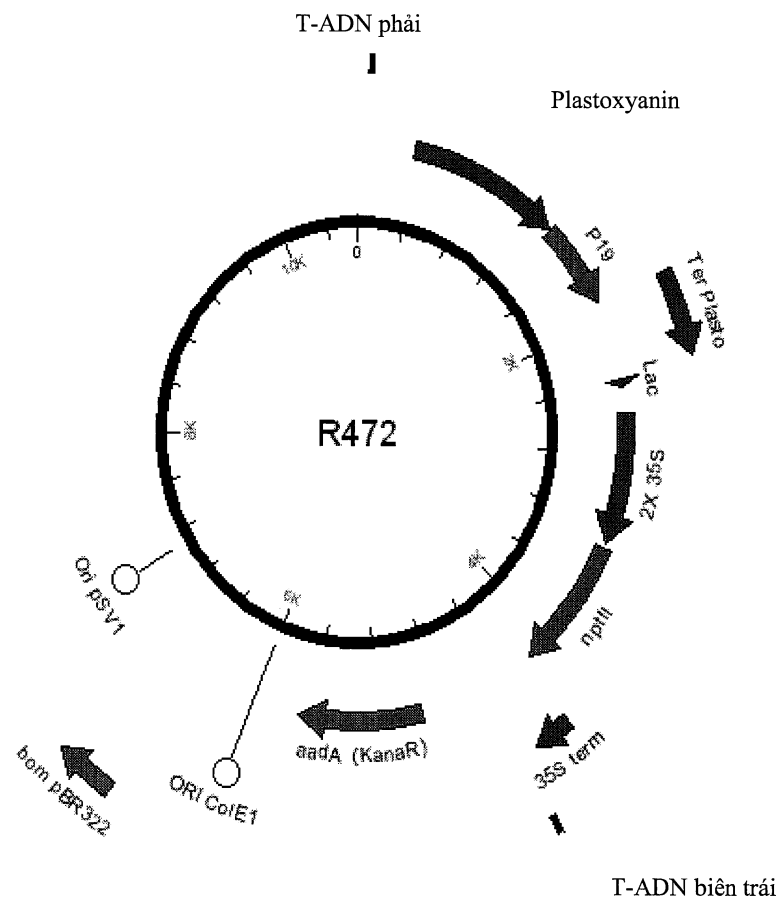
Sản phẩm thiết kế số R870, từ HindIII (trong nhiều vị trí tách dòng, ngược chiều với trình tự khởi đầu) đến EcoRI (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa HSP40 được gạch dưới in nghiêng và trình tự mã hóa HSP70 được gạch dưới.

AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGCTGGTCTGTACATTCATCTTGCCGCCCTTGCATTCACTTGGC
 CACAAAGAGTAGAGAGAAGGAAGAGAAGAGCCAGACTTCAAGAAGCGACCTTGCAAGTGCCTCGAGGGTCAGAAATGT
 ATATCATATCTATGTGAGAGAAAGGGGAACATTTGAGATGGAGTCCATTTACTTGAGGTATACTTATTATTTTATGATCAATAAAT
 TGTATACCTCTTATTTAGATCAATAAATTTGTCATTAAGCTATAATCCAAAATAAATTACGATCAAAATGCAAAATGTTAGCCAGT
 ACTTGTGTTAACTTGTATGGCATCTCTTGGTTCTTTGGCAATCACATGCCCTAAGAAATAAATAGTATCATATGATTGTGTTTGG
 TCAGACTTCAGAGTCAGATGACTCTGTTTGGATAAACAGCTTAATTAAGCGCTTATAGAATATCATATGATTGTGTTTGGTCAGACTTCAG
 ACTTCAGAGCATCTCTTGGTTCTCTGGCAATCATATGCCCTAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAG
 AGTCAGATGACCCTGTTTGGGTAACAGCTTAATTAAGTGCTTATAGAATAAGCGCTTATCATATAAGTGCTTTTGTACAGTTAT
 TTCTATGAAAGTAGAAGAAATAGTCATATTGTTTAAATATAAGTATCCTGGAGAGCTTGTGAAATAACCCAGAAAAGAACTTAT
 GGACACGTCATGAGCTGTTTACATAAGATCTCCCTAACAGTCTCAAAAGTGTTTATGCCAGTAGATAAATTCAAATAAGTCAAT
 CTAACAGACCCTAAATCCATTATGGTACCTATCATTTTAGCTTATTCCATCTTTATTAAGAATGTCATGAGATAACATAATGATA
 ACACATTATTTTGACACAAATGGGCAGATCTAGCAATTAACCTGAGAGTCTTCAAGAGTCTGTTCTTACGAAGTTCACGTC
 TCATCGGGGATTTACATTACAGAAAATATACATGACTTCAACCATACTTAAACCCCTTTTGTAAAGATAACTGAATGTTTCATATTTA
 TGCGGGCTTGACACAGAGGCAAGCTCTTTATACCTCCAGTTCACCGGGATGTTGATTTCAGAAGTCACTTGGAGAGCAATC
 CTTGTGCCAAGTTTGAAGTAATTTTGTGTAGCATATGTTGAGCTACCTACATTTACATGATCAGCTAGCATTAGCTCTTTTAC
 TTAAGTACAGAGATGAAGTTTATAGGAATGAGTATGACCATGGAGTCCGGCATGGCTTTGTAATGCCTACCTACTTTGGCCAAAC
 TCATCGGGGATTTACATTACAGAAAATATACATGACTTCAACCATACTTAAACCCCTTTTGTAAAGATAACTGAATGTTTCATATTTA
 ATGTTGGGTTGTAGTGTGTTTACTTGATTATATCCAGACAGTTACAAGTTGGACAACAAGATTGTGGGTCTGTACTGTTATTTAT
 TTATTTTTTTTTTAGCAGAAACACCTTATCTTTTGTTCGTTGAATGTAGAATGAAATAAAAGAAAGAAATATAACATCATCG
 GCCGCGCTGTCTAATTTCCGGCAGTTAGGATCCTCTCCGGTACCCGGAAGTTTCAGTAGAAGAAACAAACACCGTGACTA
 AAATGATACTATTATTTTATTATTGTGTTTTCTTTTTCTACCGGAACTTTTTGAACGGATCCCAACTCGTTCCGGGGCCGCT
 ACAACTGAAACAAAGAAGATATTTCTCTCTCTCAGAAATGTAAGTTTTCTTTACAGATACCCATTCACCATTTGATTTCAGAT
 GTGGTGACTAGAGATAAAGCATACTAATTTGACTCTTGGAAACCCATAAAGTTTTATGTTATCCGTGTTCTGGACCAATCCACTT
 GGGGGCATAACCTGTGTCTATGTGTGGTTTGGTTCCATTCTGATTTATGCGGCGACTTGTAAATTTAAATCTAGGAGGGGCA
 GACATTGAACAATCCCAATATTTTAAATACTTATGCAAGATTTTTTTTAAATGAGATGATGTGTTTGTGACTGAGATTGAGTCA
 TACATTTCACTAAGAAATGGTTCCAAGTACCAAACTATCATGACCCAGTTGCAAAACATGACGTTCCGGAGTGGTCACTTTGATA
 GTTCAATTTTCATCTTGGCTCTTATTCCCTTTATAATTTCTAATTTCTTGTGTAAACTATTTTCATGTATTATTTTCTTTAAATTT
 ACATGTCAATTTATTTGCCTCACTAATCAATTTGCATATAACAATGATAAGTGATATTTTGACTCACAAATTTACATCAATTT
 GCAGATCGTTTTATTATTTTATTGATGATTAAACAAATATAACAACTTTGCAACTAATTAACCACTGAATATAATTAACATA
 TAACTGTGAAAGTAGTTAACTCATTTTTATTTTATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAAATAA
 ACGGTATATTTTAAATAAATCTAAGCCACGTAGGAGGATAACAGGATCCCGTAGGAGGATAACATCCAATCCAACCAATCA
 CAACAATCCTGATGAGATAACCCACTTAAGCCACGCTATGTGGCAGATCTACATTATCTAAATCACACATTTCTCCACACA
 TCTGAGCCACACAAACCAATCCACATCTTTATCAGCCATCTTATAAAAAATCACACTTTGTGATGAGATTGATTCGCTT
 CAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAATGTCGGGTAAAGGAGAAGGACAGCTATC
 GGTATCGATCTTGGTACCACTTACTCTTGCCTCGGAGTATGGCAACACGACCGTGTGAGATCATTGCTAATGATCAAGGAAA
 CAGAACACGCGCTTCTACGTTGCTTTACCGACTCCGAGAGGTTGATCGGTGACGACGCTAAGAATCAGGTCCGCCATGAAC
 CCCGTTAACACCGTTTTGACGCTAAGAGGTTGATCGCTGCTGTTTCTCTGACAGCTGTTGATGAGATTGATGAGATTGATG
 GCCATTCAAGATTCAAGCCGACCTGCCGATAAGCCAATGATCTACGTCGAATACAAGGGTGAAGAGAAAGAGTTTCGACGCT
 GAGGAGATTCTTCCATGGTTCTTATTAAGATGCGTGAGATTGCTGAGGCTTACCTTGGTGTCACAATCAAGAACGCCGTTGTT
 ACCGTTCCAGCTTACTTCAAGACTCTCAGCGTCAGGCTACAAAGGATGCTGGTGTCATCGCTGGTTTGAACGTTATGCGAAT
 CATCAACGAGCTTACAGCCGCGCTATTGCCCTACGGTCTTGACAAAAAGGCTACCAGCGTTGGAGAGAAGAATGTTCTTATCT
 TCGATCTTGGTGGTGGCACCTTTGATGTCTCTCTTACCATTGAAGAGGGTATCTTTGAGGTGAAGGCAACTGCTGGTGAC
 ACCCATCTTGGTGGGGAAGATTGACAAACAGAATGGTTAACCAGTTTGCCAAGAGTTCAAGAGGAAGAGTAAGAAGGATAT
 CACCGGTAACCAAGAGCTCTTAGGAGGTTGAGAACTTCTGTGAGAGAGCGAAGAGGACTCTTTCTTCACTGCTCAGACC
 TACTCGAGATTGACATCTATACGAGGGTATCGACTTCTTCAACCATCACCCTGCTAGATTGAGGAGCTCAACATCGGA
 TCTCTTCAGGAAGTGTATGGAGCCAGTTGAGAAGTGTCTTCTGTATGCTAAGATGGACAAGAGCACTGTTTCATGATGTTGTCC
 TTGTTGGTGGTTCTACCCGATCCCTAAGGTTGACGAATGCTCCAGGACTTCTTCAACGGCAAGAGCTTTGCAAGTCTATTA
 ACCCTGATGAGGCTGTTGCCCTACGGTCTGCTGTCCAGGGAGCTATTCTCAGCGGTGAAGGAAACGAGAAGGTTCAAGATCT
 TCTATTGCTCGATCTACTCCTCTCTCCCTTGGTTTGGAACTGCCGGTGGTGTCATGACCACTTTGATCCCAAGGAACACAA
 CCATCCCAACCAAGAAGGAACAAGTCTTCTCCACCTACTCAGACAACCAACCCGGTGTGTTGATCCAGGTGTACGAAGGAGA
 GAGAGCCAGAACCAAGGACAACAACCTTCTTGGTAAATTTGAGCTCTCCGGAATTCCTCCAGCTCCTCGTGGTGTCCCCAG
 ATCAGAGTCTGCTTTGACATTGATGCAATGGTATCCTCAATGTCTGCTGAGGACAAGACCACCGGACAGAAGAACAAGAT
 CACCATCAACCAATGACAAGGGTCTCTCTCAAGGATGAGATTGAGAAGATGGTTCAAGAGGCTGAGAAGTACAAGTCCGAA
 GACGAGGAGCACAAGAAGAAGGTTGAAGCCAAGAAGCTCTCGAGAAGTACGCTTACAACATGAGGAACACCATCCAAGAGC
 AGAAGATTGGTGAGAAGCTCCCGGCTGCAGACAAGAAGAAGATCGAGGATTCTATTGAGCAGGCGATTCAATGGCTCGAGG
 GTAACCAAGTGGCTGAGGCTGATGAGTTCGAAGACAAGATGAAGGAATTGGAGAGCATCTGCAACCAATCATTGCCAAGAT
 GATAACAGGAGCTGGTGGTGAAGCCGGTGGTCCAGGTGCTCTGTTATGGACGATGATGCTCCCTCTGCTTCAGGCGGTGC
 TGGACCTAAGATCGAGGAGGTCGACTAAGAGCTCAGCTCGAATTTCCCGATCGTTCAAACATTTGGCAATAAAGTTTCTTAA
 GATTGAATCCTGTTGCCGGTCTTGCATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATG

Fig. 44b

CATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCA
 AACTAGGATAAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCGAATTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG
 TTATCCGGGGCTGGTCTGTACATTCATCTTGCCGCCCTTGCATTCACTTGGCCACAAAGAGTAGAGAGAAGGAAGAGAAGAGCCCA
 GACTTCAAGAAGCGACCTTGCAAGTGCACCTCGAGGGTCAGAACTGTATATCATATCTATGTGAGAGAAAAGGGGAACATTTGAGATG
 GAGTCCATTTACTTGAGGTATACCTTATTTTATGATCAATAAATTTGTATACCTTCTTATTTAGATCAATAAATTTGTCATTAAAGCTATAAT
 CCAAAATAAATTACGATCAAATATGCAAAATGTTAGCCAGTACTTGTGTTAAACTTGATGGCATCTCTTGGTTCTTTGGCAATCACATG
 CCTAAGAAATAAATAGTATCATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACTCTGTTTGGATAAACAGCTTAATTAAGCGC
 TTATAGAATATCATATGATTGTGTTTGGTCAGACTTCAGAGCATCTCTGTTTCTCTGGCAATCATATGCCTAAGAAATAAATAGTAT
 CATATGATTGTGTTTGGTCAGACTTCAGAGTCAGATGACCCTGTTTGGGTAAACAGCTTAATTAAGTGCTTATAGAATAAGCGCTTAT
 CATATAAGTGCTTTTGTACAGTTATTTCTATGAAAGTAGAAGAAATAGTCATATTGTTTTAATATAAGCTATCCTGGAGAGCTTGTGGA
 AATAACCAGAAAAGAACTTATGGACACGTCATGAGCTGTTTACATAAGATCTCCCTAACAGTCTCAAAAGTGTTTATGCCAGTAGATA
 AATTCAAATAAGTCAATTAACAGACCCTAAATCCATTATGGTACCTATCATTTTAGCTTATCCATCTTTATTAAGAATGTCATGAGA
 TAACATAATGATAACACATTATTTTACACAAATGGGCAGATGAGCAATTAACCTGGAAGTCTTCAAGACTTGGCCAACTGTTCTTACGAA
 GTTCACGTCCCTGAATCATGTTCTGTATGGAAGCCTGAAAGACCTCAAATTTCTAAAGGTGGCGATAAATTGAAGGTTTACAAAATA
 TACCCTGCGGGCTTGACACAGAGGCAAGCTCTTTATACCTTCCAGTTCAACGGGGATGTTGATTTTCAAGAGTCACCTTGGAGAGCAAT
 CCTTGTGCCAAGTTTGAAGTAATTTTGTGAGCATATGTTGAGCTACCTACAATTTACATGATCACCTAGCATTAGCTCTTTCACTTA
 ACTGAGAGAATGAAGTTTGTAGGAATGAGTATGACCATGGAGTCGGCATGGCTTTTGTAAATGCCCTACCCTACTTGGCCAACTGTTCTTACGG
 GGATTTACATTGAGAAAATATACATGACTTCAACCATACTTAAACCCCTTTTTGTAAGATAACTGAATGTTTCATATTTAATGTTGGGTTG
 TAGTGTTTTTACTTGATTATATCCAGACAGTTACAAGTTGGACAACAAGATTGTTGGGTCTGTACTGTTATTTATTTATTTTTTTTAGC
 AGAACACCTTATCTTTTGTTCGTTTGAATGTAGATGAAAATAAAGAAAATAACATCATCGGCCGCGCTTGTCTAATTTT
 GGGCAGTTAGGATCCTCTCGGTCACCGGAAAGTTTCAAGTAGAAGAAAACACACCGTGACTAAAATGATACTATTATTTTATTTAT
 TGTGTTTTCTTTTTCTACCGGAACTTTTAGAACGGATCCCAACTCGTTCGGGGGCGCTACAACCTGAAACAAAAGAGATATTTT
 CTCTCTCTCAGAAATGTAAGTTTTCTTTACAGATACCCATTACCATTTGATTGAGATGTGGTGACTAGAGATAAAGCATACTAATT
 TGACTCTTGAAACCCATAAAGTTTTATGTTATCCGTGTTGACCAATCCACTTGGGGGCATAACCTGTGTCTATGTGTGGTTGGT
 TTCCATTCTGATTATGCGGCGACTTGTAAATTTAAATCTAGGAGGGGCAGACATTGAACAATCCCAATATTTTAACTATGCAAG
 ATTTTTTTTATTAATGAGATGATGTGTTGTGACTGAGATTGAGTCATACATTTCACTAAGAAATGGTTCCAAGTACCAAACTATCATG
 ACCCAGTTGCAACATGACGTTCCGGGAGTGGTCACTTTGATAGTTCAATTTTCATCTTGGCTTCTTATTCCTTTTATAATTTCAATTTCTT
 CTTGTGTAAACTATTTTCATGTATTTATTTTCTTTAAATTTACATGTCATTTATTTTGGCTCACTAACTCAATTTTGCATATAACAATGAT
 AAGTGATATTTGACTCACAAAATTTACATCAAATTTGACATCGTTTTATTATGTTTCATTGGATGATTAACAAATATAACAAACTTTGCA
 ACTAATTAACCACTGAATATAATTAACATAACTGTGAAAGTAGTTAACTCATTTTTATTTTCATAGATCAAATAAGAGAAATAA
 CGGTATATTAATCCCTCCAAAAAAGGATATTTACTAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGA
 GGATAACATCCAATCCAACCAATCACAAACCTGATGAGATAACCCACTTTAAGCCACGCATCTGTGGCACATCTACATTATCTA
 AATCACACATTTCTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCATTCTATAAAAAATCACACTTTGTGAGTC
 TACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTAATTAATTAATTAATCATCTTGAGAGAAAATGTTTGGGCGCGGACC
AACAAGGAAGAGTGATAACACCAATATACGATATTCTTGGTGTTTCAAAAAGTGCTAGTGAAGATGAAATCAAGAAAGCCTATAGA
AAGGCAGCGATGAAGAACCATCCAGATAAGGGTGGGGATCCTGAGAAGTTCAAGGAGTTGGGCCAAGCATATGAAGTGTGAGCG
ATCCTGAAAAGAAAGAACTGTATGATCAATATGGTGAAGATGCCCTTAAAGAAGGAATGGGGGGAGGCGCAGGAAGCTCATTTTCAT
AATCCGTTTGATATTTTGAATCATTTTTTGGTGACGGCTTTGGTGGTGGTGGTCCCTTACGCGCAAGAAGACAGAAGCAAGGAGAA
GATGTGGTGCACTTCTATAAAGGTTTCTTGGAGGATGTGTATAACGGCACTACAAAGAAGCTATCACTTTCTAGGAATGCACTGTGC
TCAAAATGTAAAGGGAAAGGTTCAAAAAGTGGAACTGTGGAAGGTGTTTGGATGCCAGGGCACAGGTATGAAGATTACCAAGAG
GCAAAATGGACTGGGCATGATTCAACAAATGCAACACGCTGTCTGACTGCAAGGAACAGGCGAGGTCAATTAGTGAGAGAGATA
GATGCCCTCAATGCAAGGGAAACAAAGATTACTCAAGAAAAGAAAGGTGCTGGAGGTGCATGTGGAAAAGGGGATGCAGCAGGGTCA
CAAGATTGTATTCGAAGGACAAGCTGATGAAGCTCCTGATACAATCACAGGAGACATAGTTTTTGTCTTGCAAGTAAAGGGACATCC
GAAGTTTCCGAGGGAGCGTGATGACCTCCACATTGAACACAAATTTGAGCTTAACTGAGGCTCTGTGGCTTCCAGTTTAAATGTGAC
ACATCTTGATGGAAGGCAACTATTGGTCAAATCGAACCCCGGCGAAGTCATCAAGCCAGGTCAACATAAAGCTATAAATGATGAGGG
AATGCCACAACATGGTAGGCCGTTTCATGAAGGGACGCCTATACATCAAGTTTAGTGTGATTCCCGGATTCCGGGTTTTCTTCCCC
AAGCCAAAGCCTGGAATTAGAAAAGATATTACCTCAAAAGACAAGCAAGAAGCTTGTCCCAAAAGGAGGTAGATGATTGTGAGGAGAC
CACCTGTCATGATGTCAATATTGAGAGGAGATGAGTCGAAAGAAGCAACAATACCCTGAGGCATATGATGACGATGATGATGAAG
ATGATGAGCACTCGCAGCCTCGGGTGCAATGCGCTCAACAGTAGGAGCTCAGCTCGAATTTCCCGATCGTTCAACATTTGGCAA
 TAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAAC
 ATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATA
 GCGCGCAAACTAGGATAAAATATCGCGCGCGGTGTCATCTATGTTACTAGATCGAATTC

Fig. 45



Mô hình cầu disulfua của các HA tuýp A

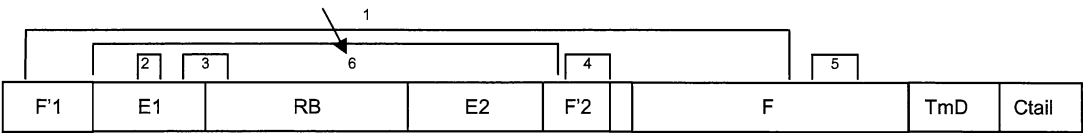


Fig. 46

Mô hình cầu disulfua của các HA typ B

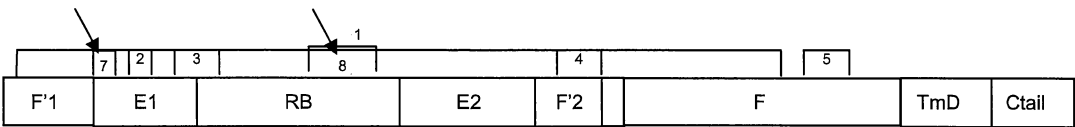


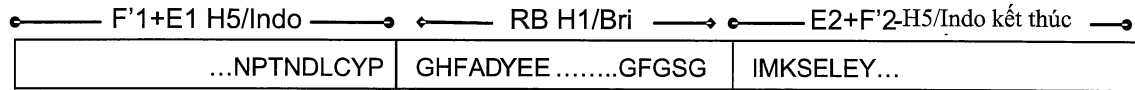
Fig. 47

Fig. 48

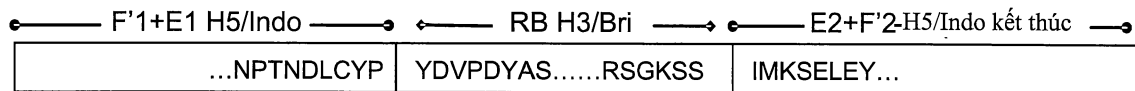
45/49

A

RB của H1/Brisbane trên H5/Indo gốc



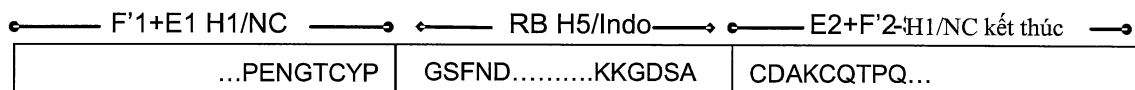
RB của H3/Brisbane trên H5/Indo gốc



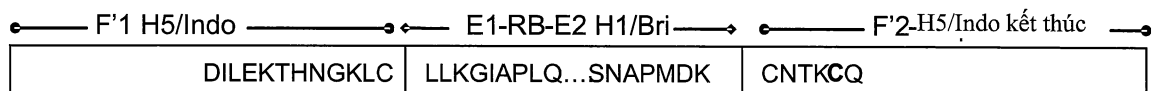
RB của B/Florida trên H5/Indo gốc



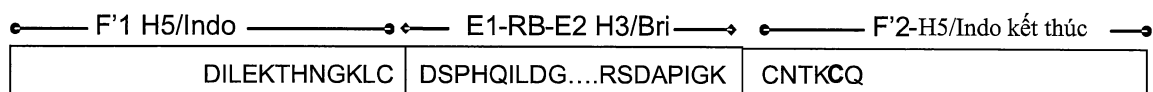
RB H5/Indo trên H1/New Cal gốc

**B**

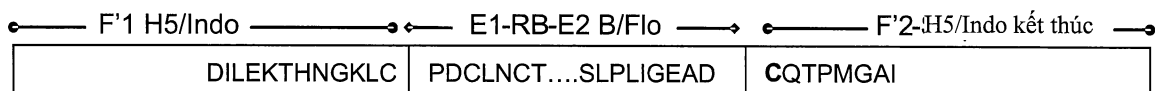
E1-RB-E2 của H1/Brisbane trên H5/Indo gốc



E1-RB-E2 của H3/Brisbane trên H5/Indo gốc



E1-RB-E2 của B/Florida trên H5/Indo gốc



E1-RB-E2 của B/Florida trên H1/New Cal gốc

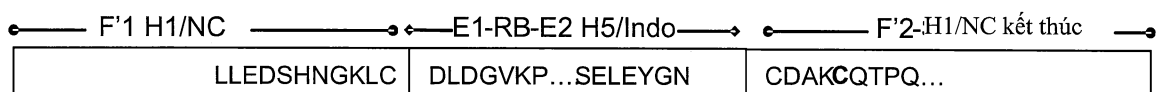


Fig. 49A

46/49

SEQ. ID. NO. 86

PDI sp - A/California/04/09

ATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCCCT
 CTCAGATCTTCGCTGACACATTATGTATAGGTTATCATGCGAACAATTCAACAGA
 CACTGTAGACACAGTACTAGAAAAGAATGTAACAGTAACACACTCTGTAAACCTT
 CTAGAAGACAAGCATAACGGGAAACTATGCAAACTAAGAGGGGTAGCCCCATTG
 CATTGGGTAAATGTAACATTGCTGGCTGGATCCTGGGAAATCCAGAGTGTGAAT
 CACTCTCCACAGCAAGCTCATGGTCCTACATTGTGGAAACACCTAGTTCAGACAA
 TGGAACGTGTTACCCAGGAGATTTTCATCGATTATGAGGAGCTAAGAGAGCAATT
 AAGCTCAGTGTTCATCATTGAAAGGTTTGAGATATTCCCAAGACAAGTTCATGG
 CCAATCATGACTCGAACAAAGGTGTAACGGCAGCATGTCCTCATGCTGGAGCA
 AAAAGCTTCTACAAAATTTAATATGGCTAGTTAAAAAAGGAAATTCATACCCAA
 AGCTCAGCAAATCCTACATTAATGATAAAGGGGAAAGAAGTCCTCGTGCTATGGG
 GCATTCACCATCCATCTACTAGTGCTGACCAACAAAGTCTCTATCAGAATGCAGA
 TACATATGTTTTGTGGGGTTCATCAAGATACAGCAAGAAGTTCAAGCCGGAAATA
 GCAATAAGACCCAAAGTGAGGGATCAAGAAGGGAGAATGAACACTATTACTGGAC
 ACTAGTAGAGCCGGGAGACAAAATAACATTGGAAGCAACTGGAAATCTAGTGGT
 ACCGAGATATGCATTGCAATGGAAAGAAATGCTGGATCTGGTATTATCATTTC
 GATACACCAGTCCACGATTGCAATACAACCTTGTCAAACACCCAAGGGTGCTATA
 AACACCAGCCTCCCATTTTCAAGATATACATCCGATCACAATTGGAAATGTCCAA
 AATATGTAAAAAGCACAAAATTGAGACTGGCCACAGGATTGAGGAATATCCCGT
 CTATTCAATCTAGAGGACTATTTGGGGCCATTGCCGGTTTCATTGAAGGGGGGTG
 GACAGGGATGGTAGATGGATGGTACGGTTATCACCATCAAAATGAGCAGGGGTG
 AGGATATGCAGCCGACCTGAAGAGCACACAGAATGCCATTGACGAGATTACTAA
 CAAAGTAAATTCTGTTATTGAAAAGATGAATACACAGTTCACAGCAGTAGGTAAA
 GAGTTCAACCACCTGGAAAAAAGAATAGAGAATTTAAATAAAAAAGTTGATGAT
 GGTTCCTGGACATTTGGACTTACAATGCCGAACCTGTTGGTTCTATTGGAAATG
 AAAGAACTTTGGACTACCACGATTCAAATGTGAAGAACTTATATGAAAAGGTAAG
 AAGCCAGCTAAAAACAATGCCAAGGAAATTGGAAACGGCTGCTTTGAATTTTA
 CCACAAATGCGATAACACGTGCATGGAAAGTGTCAAAAATGGGACTTATGACTA
 CCAAAAATACTCAGAGGAAGCAAAATTAACAGAGAAGAAATAGATGGGGTAAA
 GCTGGAATCAACAAGGATTTACCAGATTTTGGCGATCTATTCAACTGTCGCCAGT
 TCATTGGTACTGGTAGTCTCCCTGGGGCAATCAGTTTCTGGATGTGCTCTAATG
 GGTCTCTACAGTGTAGAATATGTATT

Fig. 49B

SEQ. ID. NO. 87

PDI sp - A/California/04/09

MAKNVAIFGLLFSLVLVPSQIFADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLLEDKHNGKLCKLRG
 VAPLHLGKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELREQLSSVSSFERFEI
 FPKTSSWPNHDSNKGVTAAACPHAGAKSFYKNLIWLKKGNSYPKLSKSYINDKGKEVLVLWGIHHPST
 SADQQSLYQNADTYVFGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTLVEPGDKITFEATGNLWVPR
 YAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINSLPFQNIHPITIGKCPKYVKSTKLRLATGLRNIPSI
 QSRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFT
 AVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENERLTDYHDSNVKNLYEKVRSQKNNAKEI
 GNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTRIYQILAIYSTVASSLVLVVSL
 GAISFWMCNSGLQCRICI

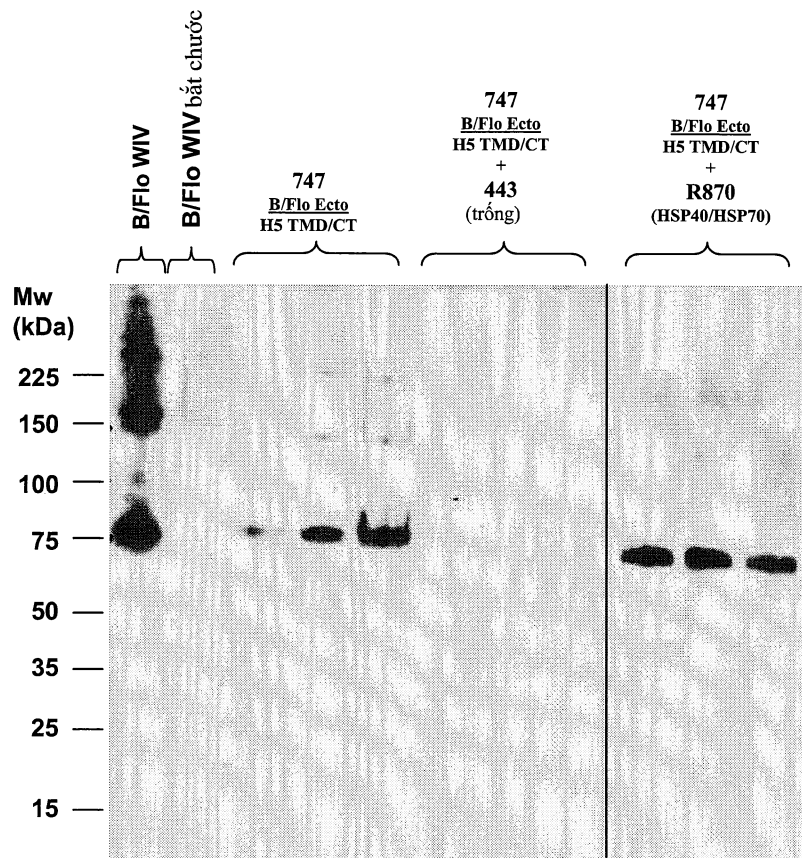


Fig. 50

SEQ. ID. NO. 88

Trình tự khởi đầu 2X35S

GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAA
GATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGG
GTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACT
TTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATC
ATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTG
GTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAA
GACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATG
GTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCT
CAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCG
GAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAA
GATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAA
AGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAG
ATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
ACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAA
GGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATAT
AAGGAAGTTCATTTCAATTTGGAGAGG

Fig. 51A

SEQ. ID. SỐ. 93

Sản phẩm thiết kế 747, từ PacI (ngược chiều với trình tự khởi đầu 35S) đến AscI (ngay cùng chiều với nhân tố kết thúc NOS). Trình tự mã hóa HA thể khảm được gạch dưới. Trình tự khởi đầu 2X35S in đậm.

TTAATTAAGTCGACAAGCTTGCATGCCTGCAGGTCAACATGGTGGAGCACGACACACTTGTCTACTC
CAAAAATATCAAAGATACAGTCTCAGAAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATC
CGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAG
GTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAG
TGGTCCCAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCT
TCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAA
AGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCC
TCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTAC
AAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAG
ATGGACCCCAACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGT
GGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCCTT
CCTCTATATAAGGAAGTTCAATTTCAATTTGGAGAGGTATTAATAATCTTAATAGGTTTTGATAAAAGCGAAC
GTGGGGAACCCGAACCAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTCTTG
TCTTCTTGCGTGAGCGATCTTCAACGTTGTGATAGTGTGCTTGGCACCAGTACAACGTTTTCTTCTC
ACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCT
CTATTCTTGTGGTGTGGTCTTGGGAAAAGAAAGCTTGTGGAGGCTGCTGTTCAAGCCCATACATTA
CTTGTTACGATTCTGCTGACTTTGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGC
CTGTACTTCTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAT
TTTCCCGTGGTTTTTGAACCTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTGGGGC
CCATGGCGAAAAACGTTGCGATTTTGGGCTTATTGTTTTCTTCTTGTGTTGGTTCCCTTCTCAGATCT
TCGCTGATCGAATCTGCACTGGAATAACATCTTCAAACCTCACCTCATGTGGTCAAAACAGCCACTCAA
GGGGAGGTCAATGTGACTGGTGTGATACCACTAACAAACACCAACAAAAATCTTATTTTGCAAATCT
CAAAGGAACAAGGACCAGAGGGAAACTATGCCAGACTGTCTCAACTGCACAGATCTGGATGTGGCT
TTGGGCAGACCAATGTGTGTGGGGACCAACCTTGGCGGAAGGCTTCAATACTCCACGAAGTCAAAC
CTGTTACATCCGGGTGCTTTCTATAATGCACGACAGAAACAAAAATCAGGCAACTACCAATCTTCTC
AGAGGATATGAAAAATACAGGCTATCAACCCCAAAACGTCATCGATGCGGAAAAGGCCAGGAGGAC
CCTACAGACTTGGAACTCAGGATCTTGCCCTAACGCTACCAGTAAGAGCGGATTTTTCGCAACAATG
GCTTGGGCTGTCCCAAAGGACAACAACAAAAATGCAACGAACCCACTAACAGTAGAAGTACCATACAT
TTGTACAGAAGGGGAAGACCAATCACTGTTTGGGGTTCATTGAGATAACAAAAACCAAAATGAAGA
ACCTCTATGGAGACTCAAATCCTCAAAGTTTCACTCTGCTAATGGAGTAACCAACACACTATGTTT
CTCAGATTGGCAGCTTCCAGATCAAACAGAAGACGGAGGACTACCACAAAGCGGCAGGATTGTTGT
TGATTACATGATGCAAAAACCTGGGAAAAAGGAACAATTGTCTACCAAGAGGTGTTTTGTTGCCTC
AAAAGGTGTGGTGGCGAGTGGCAGGAGCAAGTAATAAAAGGGTCCCTTGCCTTTAATTGGTGAAGC
AGATTGCCTTCATGAAAAATACGGTGGATTAAACAAAAGCAAGCCTTACTACACAGGAGAAACATGCAA
AAGCCATAGGAAATGCCCCAATATGGGTGAAAACACCTTTGAAGCTCGCCAATGGAACCAATATAGA
CCTCCTGCAAAACTATTAAAGGAAAGGGGTTTCTTGGAGCTATTGCTGGTTTCTAGAAAGGAGGATG
GGAAGGAATGATTGCAGGCTGGCAGGATACACATCTCACGGAGCACATGGAGTGGCAGTGGCGGC
GGACCTTAAGAGTACGCAAGAAGCTATAAACAAGATAACAAAAATCTCAATTCTTTGAGTGAGCTAGA
AGTAAAGAATCTTCAAAGACTAAGTGGTGGCATGGATGAAGTCCACAACGAAATACTCGAGCTGGATG
AGAAAGTGGATGATCTCAGAGCTGACACTATAAGCTCGCAATAGAACTTGCAGTCTTGTCTTCCAAC
GAAGGAATAATAAACAGTGAAGATGAGCATCTATTGGCACTTGAGAGAAAACTAAAGAAAAATGCTGGG
TCCCTCTGCTGTAGAGATAGGAAATGGATGCTTCGAAACCAAAACACAAGTGAACCAAGACCTGCTTAG
ACAGGATAGCTGCTGGCACCTTTAATGCAGGAGAATTTCTCTCCCACTTTTGATTCACTGAACATTA
CTGCTGCATCTTTAAATGATGATGGATTGGATAACTACCAAAATACTGTCAATTTATTCAACAGTGGCGA
GTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAA
TGCAGAATTTGCATTTAAAGGCCTATTTCTTTAGTTTGAATTTACTGTTATTCCGGTGTGCATTTCTATG
TTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGT
TTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTTATTAATAAAGTTTCTTAAGA
CCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGA
TTGAATCCTGTTGCCGGTCTTGCATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAA
TTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTA
ATACGCGATAGAAAACAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGT
TACTAGATTCTAGAGTCTCAAGCTT**GGCGCGCC**

Fig. 51B

DANH MỤC TRÌNH TỰ

- <110> Medicago Inc.
COUTURE, MANON
DARGIS, MICHELE
LAVOIE, PIERRE-OLIVIER
VEZINA, LOUIS-PHILIPPE
DAOUST, MARC-ANDRE
- <120> PHƯƠNG PHÁP SẢN XUẤT HẠT TƯƠNG TỰ VIRUT CÚM THỂ KHẢM Ở THỰC VẬT VÀ AXIT NUCLEIC MÃ HÓA POLYPEPTIT CỦA HẠT TƯƠNG TỰ VIRUT NÀY
- <130> V82197WO
- <140> không được chỉ định
<141> không được chỉ định
- <150> 61/220,161
<151> 2009-06-24
- <160> 113
- <170> PatentIn version 3.4
- <210> 1
<211> 32
<212> ADN
<213> Trình tự nhân tạo
- <220>
<223> XmaI-pPlas.c tổng hợp
- <400> 1
agttcccccgg gctggtatat ttatatgttg tc 32
- <210> 2
<211> 46
<212> ADN
<213> Trình tự nhân tạo
- <220>
<223> SacI-ATG-pPlas.r tổng hợp
- <400> 2
aatagagctc cattttctct caagatgatt aattaattaa ttagtc 46
- <210> 3
<211> 46
<212> ADN
<213> Trình tự nhân tạo
- <220>
<223> SacI-PlasTer.c tổng hợp
- <400> 3
aatagagctc gttaaaatgc ttcttcgtct cctatttata atatgg 46
- <210> 4
<211> 48
<212> ADN

<213> Trình tự nhân tạo
 <220>
 <223> EcoRI-PlasTer.r tổng hợp
 <400> 4
 ttacgaattc tccttcctaa ttggtgtact atcatttatc aaagggga 48
 <210> 5
 <211> 25
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> Plasto-443C tổng hợp
 <400> 5
 gtattagtaa ttagaatttg gtgtc 25
 <210> 6
 <211> 44
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> SpHa(Ind)-Plasto-r
 <400> 6
 gcaagaagaa gcactatddd ctccattttc tctcaagatg atta 44
 <210> 7
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> Plasto-SpHa.c
 <400> 7
 ttaatcatct tgagagaaaa tggagaaaat agtgcttctt cttgc 45
 <210> 8
 <211> 38
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> HA(Ind)-Sac.r
 <400> 8
 actttgagct cttaaagca aattctgcat tgtaacga 38
 <210> 9
 <211> 24
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> pBinPlus.2613c tổng hợp
 <400> 9

aggaagggaa gaaagcgaaa ggag 24

<210> 10
 <211> 56
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Mut-ATG115.r tổng hợp

<400> 10
 gtgccgaagc acgatctgac aacgttgaag atcgctcacg caagaaagac aagaga 56

<210> 11
 <211> 52
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Mut-ATG161.c tổng hợp

<400> 11
 gttgtcagat cgtgcttcgg caccagtaca acgttttctt tcactgaagc ga 52

<210> 12
 <211> 25
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> LC-C5-1.110r tổng hợp

<400> 12
 tctcctggag tcacagacag ggtgg 25

<210> 13
 <211> 52
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> E1 H1B-E1 H5I.r tổng hợp

<400> 13
 tcatagtcag cgaaatgccc tgggtaacag aggtcattgg ttggattggc ct 52

<210> 14
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> E1 H5N-E1 H1B.c tổng hợp

<400> 14
 atgacctctg ttacccaggg catttcgctg actatgagga actgaggg 48

<210> 15
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> E2 H5I-RB H1B.r tổng hợp

 <400> 15
 ccaattcact tttcataatt cctgatccaa agcctctact cagtgcga 48

 <210> 16
 <211> 51
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> RB H1B-E2 H5I.c tổng hợp

 <400> 16
 ggcttttgat caggaattat gaaaagtga ttggaatatg gtaactgcaa c 51

 <210> 17
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> E1 H1B-F(prime)1 H5I.r tổng hợp

 <400> 17
 ggctattcct tttaataggc agagcttccc gttgtgtgtc ttttccag 48

 <210> 18
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> F'1 H5N-E1 H1B.c tổng hợp

 <400> 18
 aacgggaagc tctgcctatt aaaaggaata gcccactac aattgggt 48

 <210> 19
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> F(gốc)2 H5I-E2 H1B.r tổng hợp

 <400> 19
 ggagtttgac acttggtgtt gcatttatcc attggtgcat ttgagttg 48

 <210> 20
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> E2 H1B-F(prime)2 H5I.c tổng hợp

 <400> 20
 aatgcaccaa tggataaatg caacaccaag tgtcaaactc caatgggg 48

<210> 21
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> E1 H5I-E1 H1NC.r tổng hợp

 <400> 21
 tcttcatagt cgttgaaact ccctgggtaa catgttccat tctcagga 48

 <210> 22
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> E1 H1NC-E1 H5I.c tổng hợp

 <400> 22
 ctgagaatgg aacatgttac ccagggagtt tcaacgacta tgaagaac 48

 <210> 23
 <211> 51
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> E2 H1NC-RB H5I.r tổng hợp

 <400> 23
 atttgaggtg atgattgctg agtccccctt cttgacaatt ttgtatgcat a 51

 <210> 24
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> RB H5I-E2 H1NC.c tổng hợp

 <400> 24
 gtcaagaaag gggactcagc aatcatcacc tcaaatgcac caatggat 48

 <210> 25
 <211> 43
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Tổng hợp HA-SacI.r

 <400> 25
 ttaacttaga gctcttagat gcatattcta cactgcaaag acc 43

 <210> 26
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ApaI-H1B.c tổng hợp

 <400> 26
 tgtcgggccc atgaaagtaa aactactggt cctgttatgc acatt 45

 <210> 27
 <211> 46
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> StuI-H1B.r tổng hợp

 <400> 27
 aaataggcct ttagatgcat attctacact gtaaagaccc attgga 46

 <210> 28
 <211> 49
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> SpPDI-H1B.c tổng hợp

 <400> 28
 ttctcagatc ttcgctgaca caatatgtat aggctaccat gctaacaac 49

 <210> 29
 <211> 47
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> SacI-H1B.r tổng hợp

 <400> 29
 cttagagctc ttagatgcat attctacact gtaaagaccc attggaa 47

 <210> 30
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ApaI-SpPDI.c tổng hợp

 <400> 30
 ttgtcgggcc catggcgaaa aacgttgcca ttttcggctt attgt 45

 <210> 31
 <211> 39
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ApaI-H5 (A-Indo).1c tổng hợp

 <400> 31
 tgtcgggccc atggagaaaa tagtgcttct tcttgcaat 39

<210> 32
 <211> 37
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> H5 (A-Indo)-StuI.1707r tổng hợp

 <400> 32
 aaataggcct ttaaagcaa attctgcatt gtaacga 37

 <210> 33
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> H3B-SpPDI.r tổng hợp

 <400> 33
 tgtcatttcc gggaagtttt tgagcgaaga tctgagaagg aacca 45

 <210> 34
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> SpPDI-H3B.c tổng hợp

 <400> 34
 tctcagatct tcgctcaaaa acttcccga aatgacaaca gcacg 45

 <210> 35
 <211> 44
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> StuI-H3B.r tổng hợp

 <400> 35
 aaaataggcc ttcaaagca aatgttgac ctaatgttg cttt 44

 <210> 36
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> TmD H5I-H3B.r tổng hợp

 <400> 36
 atttggttaag ttcctattga cttcagctca acgcccttga tctgg 45

 <210> 37
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

 <220>

<223> H3B-TmD H5I.c tổng hợp
 <400> 37
 tgagctgaag tcaataggaa cttaccaaact actgtcaatt tattcaac 48
 <210> 38
 <211> 50
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> HBF-SpPDI.r tổng hợp
 <400> 38
 gttattccag tgcagattcg atcagcgaag atctgagaag gaaccaacac 50
 <210> 39
 <211> 50
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> SpPDI-HBF.c tổng hợp
 <400> 39
 cagatcttcg ctgactgaat ctgcactgga ataacatctt caaactcacc 50
 <210> 40
 <211> 46
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> StuI-HBF.r tổng hợp
 <400> 40
 aaaataggcc tttatagaca gatggagcat gaaacgttgt ctctgg 46
 <210> 41
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> TmD H5I-B Flo.r tổng hợp
 <400> 41
 tgacagtatt tggtagttat ccaatccatc atcatttaaa gatgc 45
 <210> 42
 <211> 50
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> B Flo-TmD H5I.c tổng hợp
 <400> 42
 ggattggata actaccaaact actgtcaatt tattcaacag tggcgagttc 50
 <210> 43

<211> 20
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Hsp-40Luz.1c tổng hợp

 <400> 43
 atgtttgggc gcggaccaac 20

 <210> 44
 <211> 31
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Hsp40Luz-SacI.1272r tổng hợp

 <400> 44
 agctgagctc ctactgttga gcgcattgca c 31

 <210> 45
 <211> 36
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Hsp40Luz-Plasto.r tổng hợp

 <400> 45
 gttgggtccgc gcccaaacat tttctctcaa gatgat 36

 <210> 46
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Hsp70Ara.1c tổng hợp

 <400> 46
 atgtcgggta aaggagaagg a 21

 <210> 47
 <211> 33
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Hsp70Ara-SacI.1956r tổng hợp

 <400> 47
 agctgagctc ttagtcgacc tcctcgatct tag 33

 <210> 48
 <211> 37
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Hsp70Ara-Plasto.r tổng hợp

<400> 48	
tcctttctcct ttacccgaca ttttctctca agatgat	37
<210> 49	
<211> 34	
<212> ADN	
<213> Trình tự nhân tạo	
<220>	
<223> supP19-plasto.r tổng hợp	
<400> 49	
ccttgatatag ctcgttccat tttctctcaa gatg	34
<210> 50	
<211> 20	
<212> ADN	
<213> Trình tự nhân tạo	
<220>	
<223> supP19-1c tổng hợp	
<400> 50	
atggaacgag ctatacaagg	20
<210> 51	
<211> 32	
<212> ADN	
<213> Trình tự nhân tạo	
<220>	
<223> SupP19-SacI.r tổng hợp	
<400> 51	
agtcgagctc ttactcgctt tcttttttcga ag	32
<210> 52	
<211> 1719	
<212> ADN	
<213> Trình tự nhân tạo	
<220>	
<223> B-Plasto-SP gốc-H5 A/Indonesia/5/05 tổng hợp	
<400> 52	
aagcttatgg agaaaatagt gcttcttctt gcaatagtca gtcttgtaa aagtgatcag	60
atttgcattg gttacatgc aaacaattca acagagcagg ttgacacaat catggaaaag	120
aacgttactg ttacacatgc ccaagacata ctggaaaaga cacacaacgg gaagctctgc	180
gatctagatg gagtgaagcc tctaatttta agagattgta gtgtagctgg atggctcctc	240
gggaacccaa tgtgtgacga attcatcaat gtaccggaat ggtcttacat agtgagagaag	300
gccaatccaa ccaatgacct ctgttaccca gggagtttca acgactatga agaactgaaa	360
cacctattga gcagaataaa ccattttgag aaaattcaaa tcatccccaa aagttcttgg	420
tccgatcatg aagcctcatc aggagttagc tcagcatgtc catacctggg aagtcacctc	480

ttttttagaa atgtggtatg gcttatcaaa aagaacagta catacccaac aataaagaaa 540
 agctacaata ataccaacca agaggatctt ttggtactgt ggggaattca ccatcctaata 600
 gatgcggcag agcagacaag gctatatcaa aaccaacca cctatatattc cattgggaca 660
 tcaacactaa accagagatt ggtacaaaa atagctacta gatccaaagt aaacgggcaa 720
 agtggaagga tggagttctt ctggacaatt ttaaaaccta atgatgcaat caacttcgag 780
 agtaatggaa atttcattgc tccagaatat gcatacaaaa ttgtcaagaa aggggactca 840
 gcaattatga aaagtgaatt ggaatatggt aactgcaaca ccaagtgtca aactccaatg 900
 ggggcgataa actctagtat gccattccac aacatacacc ctctcacat cggggaatgc 960
 cccaaatatg tgaaatcaaa cagattagtc cttgcaacag ggctcagaaa tagccctcaa 1020
 agagagagca gaagaaaaaa gagaggacta tttggagcta tagcaggttt tatagaggga 1080
 ggatggcagg gaatggtaga tggttggtat ggggtaccacc atagcaatga gcaggggagt 1140
 ggggtacgctg cagacaaaga atccactcaa aaggcaatag atggagtcac caataaggtc 1200
 aactcaatca ttgacaaaat gaacactcag tttgaggccg ttggaaggga atttaataac 1260
 ttagaaagga gaatagagaa tttaaacaag aagatggaag acgggtttct agatgtctgg 1320
 acttataatg ccgaacttct gggtctcatg gaaaatgaga gaactctaga ctttcatgac 1380
 tcaaattgta agaacctcta cgacaaggtc cgactacagc ttagggataa tgcaaaggag 1440
 ctgggtaacg gttgtttcga gttctatcac aaatgtgata atgaatgtat ggaaagtata 1500
 agaaacggaa cgtacaacta tccgcagtat tcagaagaag caagattaaa aagagaggaa 1560
 ataagtggg taaaattgga atcaatagga acttaccaaa tactgtcaat ttattcaaca 1620
 gtggcgagtt ccctagcact ggcaatcatg atggctggtc tatctttatg gatgtgctcc 1680
 aatggatcgt tacaatgcag aatttgcatt taagagctc 1719

<210> 53

<211> 3194

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Catxet biểu hiện 660 tổng hợp từ HindIII đến Eco-R1

<400> 53

aagcttgcta gcggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta 60
 tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagtttagca agtgtgtaca 120
 tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag 180
 taaagaaata tggatgataa gaacaagagt agtgatattt tgacaacaat tttgttgcaa 240
 catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa 300

aggaagaggg	agaataaaaa	cataatgtga	gtatgagaga	gaaagttgta	caaaagttgt	360
acaaaaatag	ttgtacaaat	atcattgagg	aatttgacaa	aagctacaca	aataaggggt	420
aattgctgta	aataaataag	gatgacgcat	tagagagatg	taccattaga	gaatttttgg	480
caagtcatta	aaaagaaaga	ataaattatt	tttaaaatta	aaagttgagt	catttgatta	540
aacatgtgat	tatttaatga	attgatgaaa	gagttggatt	aaagttgtat	tagtaattag	600
aatttggtgt	caaatttaat	ttgacatttg	atcttttcct	atatattgcc	ccatagagtc	660
agttaactca	tttttatatt	tcatagatca	aataagagaa	ataacggtat	attaatccct	720
ccaaaaaaaa	aaaacggtat	atttactaaa	aatctaagc	cacgtaggag	gataacagga	780
tccccgtagg	aggataacat	ccaatccaac	caatcacac	aatcctgatg	agataacca	840
ctttaagccc	acgcatctgt	ggcacatcta	cattatctaa	atcacacatt	cttcacaca	900
tctgagccac	acaaaaacca	atccacatct	ttatcaccca	ttctataaaa	aatcacactt	960
tgtgagtcta	cactttgatt	cccttcaaac	acatacaaag	agaagagact	aattaattaa	1020
ttaatcatct	tgagagaaaa	tgagagaaaat	agtgttctt	cttgcaatag	tcagtcttgt	1080
taaaagtgat	cagatttgca	ttggttacca	tgcaaacaat	tcaacagagc	aggttgacac	1140
aatcatggaa	aagaacgtta	ctgttacaca	tgcccaagac	atactggaaa	agacacacaa	1200
cgggaagctc	tgcatctag	atggagtga	gcctctaatt	ttaagagatt	gtagtgtagc	1260
tggatggctc	ctcgggaacc	caatgtgtga	cgaattcatc	aatgtaccgg	aatgggtctta	1320
catagtggag	aaggccaatc	caaccaatga	cctctgttac	ccaggagatt	tcaacgacta	1380
tgaagaactg	aaacacctat	tgagcagaat	aaaccatttt	gagaaaattc	aatcatccc	1440
caaaagttct	tgttccgatc	atgaagcctc	atcaggagtt	agctcagcat	gtccatacct	1500
gggaagtccc	tcctttttta	gaaatgtggt	atggcttata	aaaaagaaca	gtacataccc	1560
aacaataaag	aaaagctaca	ataataccaa	ccaagaggat	cttttggtac	tgtggggaat	1620
tcaccatcct	aatgatgcgg	cagagcagac	aaggctatat	caaaacccaa	ccacctatat	1680
ttccattggg	acatcaacac	taaaccagag	attggtacca	aaaatagcta	ctagatccaa	1740
agtaaacggg	caaagtggaa	ggatggagtt	cttctggaca	attttaaaac	ctaattgatgc	1800
aatcaacttc	gagagtaatg	gaaatttcat	tgctccagaa	tatgcataca	aaattgtcaa	1860
gaaaggggac	tcagcaatta	tgaaaagtga	attggaatat	ggtaactgca	acaccaagtg	1920
tcaaactcca	atgggggcga	taaactctag	tatgccattc	cacaacatac	accctctcac	1980
catcggggaa	tgccccaat	atgtgaaatc	aaacagatta	gtccttgcaa	cagggtcag	2040
aaatagccct	caaagagaga	gcagaagaaa	aaagagagga	ctatttggag	ctatagcagg	2100

```

ttttatagag ggaggatggc agggaatggg agatgggttg tatgggtacc accatagcaa 2160
tgagcagggg agtgggtacg ctgcagacaa agaatccact caaaaggcaa tagatggagt 2220
caccaataag gtcaactcaa tcattgacaa aatgaacact cagtttgagg ccgttggaag 2280
ggaatttaat aacttagaaa ggagaataga gaatttaaac aagaagatgg aagacggggt 2340
tctagatgtc tggacttata atgccgaact tctggttctc atggaaaatg agagaactct 2400
agactttcat gactcaaatg ttaagaacct ctacgacaag gtccgactac agcttaggga 2460
taatgcaaag gagctgggta acggttggtt cgagttctat cacaaatgtg ataatgaatg 2520
tatggaaagt ataagaaacg gaacgtacaa ctatccgcag tattcagaag aagcaagatt 2580
aaaaagagag gaaataagtg gggtaaaatt ggaatcaata ggaacttacc aaatactgtc 2640
aattttattca acagtggcga gttccctagc actggcaatc atgatggctg gtctatcttt 2700
atggatgtgc tccaatggat cgttacaatg cagaatttgc atttaagagc tctaagttaa 2760
aatgcttctt cgtctcctat ttataatatg gtttggtatt gttaattttg ttcttgtaga 2820
agagcttaat taatcgttgt tgttatgaaa tactatttgt atgagatgaa ctggtgtaat 2880
gtaattcatt tacataagtg gagtcagaat cagaatgttt cctccataac taactagaca 2940
tgaagacctg ccggtacaa ttgtcttata tttgaacaac taaaattgaa catcttttgc 3000
cacaacttta taagtgggta atatagctca aatatatggg caagttcaat agattaataa 3060
tggaatatatc agttatcgaa attcattaac aatcaactta acgttattaa ctactaattt 3120
tatatcatcc cctttgataa atgatagtag accaattagg aaggagcatg ctcgaggcct 3180
ggctggccga attc 3194

```

<210> 54

<211> 1711

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự mã hóa H1/NC loại đại tổng hợp thiếu TmD và Ctail

<400> 54

```

atgaaagcaa aactactggg cctggttatgt acatttacag ctacatatgc agacacaata 60
tgtatagggt accatgccaa caactcaacc gacactgttg acacagtact tgagaagaat 120
gtgacagtga cacactctgt caacctactt gaggacagtc acaatggaaa actatgtcta 180
ctaaaaggaa tagccccact acaattgggt aattgcagcg ttgccgatg gatcttagga 240
aaccagaat gcgaattact gatttccaag gaatcatggg cctacattgt agaaacacca 300
aatcctgaga atggaacatg ttaccagggt tatttcgccg actatgagga actgaggagg 360
caattgagtt cagtatcttc atttgagaga ttcgaaatat tccccaaaga aagctcatgg 420

```

cccaaccaca ccgtaaccgg agtatcagca tcatgctccc ataatgggaa aagcagtttt 480
 tacagaaatt tgctatggct gacggggaag aatggtttgt acccaaacct gagcaagtcc 540
 tatgtaaaca acaaagagaa agaagtcctt gtactatggg gtgttcatca cccgcctaac 600
 ataggggaacc aaagggccct ctatcataca gaaaatgctt atgtctctgt agtgtcttca 660
 cattatagca gaagattcac cccagaaata gccaaaagac ccaaagtaag agatcaggaa 720
 ggaagaatca actactactg gactctgctg gaacctgggg atacaataat atttgaggca 780
 aatggaaatc taatagcgcc atggatatgct tttgcactga gtagaggctt tggatcagga 840
 atcatcacct caaatgcacc aatggatgaa tgtgatgcga agtgtcaaac acctcaggga 900
 gctataaaca gcagtcttcc tttccagaat gtacaccag tcacaatagg agagtgtcca 960
 aagtatgtca ggagtgcaaa attaaggatg gttacaggac taaggaacat cccatccatt 1020
 caatccagag gtttgtttg agccattgcc ggtttcattg aaggggggtg gactggaatg 1080
 gtagatgggt ggtatggtta tcatcatcag aatgagcaag gatctggcta tgctgcagat 1140
 caaaaaagta cacaaatgc cattaacggg attacaaaca aggtgaattc tgtaattgag 1200
 aaaatgaaca ctcaattcac agctgtgggc aaagaattca acaaattgga aagaaggatg 1260
 gaaaacttaa ataaaaaagt tgatgatggg tttctagaca tttggacata taatgcagaa 1320
 ttgttggttc tactggaaaa tgaaaggact ttggatttcc atgactccaa tgtgaagaat 1380
 ctgtatgaga aagtaaaaag ccaattaaag aataatgcca aagaaatagg aaacgggtgt 1440
 tttgaattct atcacaagtg taacaatgaa tgcattggaga gtgtgaaaaa tggaaacttat 1500
 gactatccaa aatattccga agaatcaaag ttaaacaggg agaaaattga tggagtgaaa 1560
 ttggaatcaa tgggagtcta tcagattctg gcgatctact caactgtcgc cagttccctg 1620
 gttcttttgg tctccctggg ggcaatcagc ttctggatgt gttccaatgg gtctttgcag 1680
 tgtagaatat gcatctgaga ccagaatttc a 1711

<210> 55

<211> 1556

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> BglII-H1 A/NewCaledonia/20/99-SacI/StuI tổng hợp.

<400> 55

agatcttcgc tgacacaata tgtataggct accatgcca caactcaacc gacactgttg 60
 acacagtact tgagaagaat gtgacagtga cacactctgt caacctactt gaggacagtc 120
 acaatggaaa actatgtcta ctaaaaggaa tagccccact acaattgggt aattgcagcg 180
 ttgccggatg gatcttagga aaccagaat gcgaattact gatttccaag gaatcatggt 240

cctacattgt agaaacacca aatcctgaga atggaacatg ttacccaggg tatttcgccg 300
 actatgagga actgagggag caattgagtt cagtatcttc atttgagaga ttcgaaatat 360
 tccccaaaga aagctcatgg cccaaccaca ccgtaaccgg agtatcagca tcatgctccc 420
 ataatgggaa aagcagtttt tacagaaatt tgctatggct gacggggaag aatggtttgt 480
 acccaaacct gagcaagtcc tatgtaaaca acaaagagaa agaagtcctt gtactatggg 540
 gtgttcatca cccgcctaac atagggaacc aaagggcact ctatcataca gaaaatgctt 600
 atgtctctgt agtgtcttca cattatagca gaagattcac ccagaaaata gccaaaagac 660
 ccaaagtaag agatcaggaa ggaagaatca actactactg gactctgctg gaacctgggg 720
 atacaataat atttgaggca aatggaaatc taatagcgcc atgggatgct tttgcactga 780
 gtagaggctt tggatcagga atcatcacct caaatgcacc aatggatgaa tgtgatgcga 840
 agtgtcaaac acctcaggga gctataaaca gcagtcttcc tttccagaat gtacaccag 900
 tcacaatagg agagtgtcca aagtatgtca ggagtgcaa attaaggatg gttacaggac 960
 taaggaacat cccatccatt caatccagag gtttgtttg agccattgcc ggtttcattg 1020
 aaggggggtg gactggaatg gtagatgggt ggtatgggtta tcatcatcag aatgagcaag 1080
 gatctggcta tgctgcagat caaaaaagta cacaaaatgc cattaacggg attacaaaca 1140
 aggtcaattc tgtaattgag aaaatgaaca ctcaattcac agctgtgggc aaagagttca 1200
 acaaattgga aagaaggatg gaaaacttaa ataaaaaagt tgatgatggg tttctagaca 1260
 tttggacata taatgcagaa ttgttggttc tactggaaaa tgaaaggact ttggatttcc 1320
 atgactccaa tgtgaagaat ctgtatgaga aagtaaaaag ccaattaaag aataatgcca 1380
 aagaaatagg aaacgggtgt tttgagttct atcacaagtg taacaatgaa tgcatggaga 1440
 gtgtgaaaaa tggtagctat gactatccaa aatattccga agaatcaaag ttaaacaggg 1500
 agaaaattga tggagtgaaa ttggaatcaa tgggagtata ctaagagctc aggcct 1556

 <210> 56
 <211> 219
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> KpnI-H1 A/NewCaledonia/20/99 TmD+Ctail-SacI/StuI tổng hợp.

 <400> 56
 ggtacctatg actatccaaa atattccgaa gaatcaaagt taaacaggga gaaaattgat 60
 ggagtgaaat tggaatcaat gggagtatac cagattctgg cgatctactc aactgtcgcc 120
 agttccctgg ttcttttgggt ctccctgggg gcaatcagct tctggatgtg ttccaatggg 180
 tctttgcagt gtagaatatg catctaagag ctcaggcct 219

<210> 57
 <211> 1781
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Peptit báo hiệu của gen protein disulfua isomeraza (PDI) ở cò linh
 lăng tổng hợp

<400> 57

```

ccaaatcctt aacattcttt caacaccaac aatggcgaaa aacgttgcca ttttcggttt      60
attgttttct cttcttctgt tggttccttc tcagatcttc gctgaggaat catcaactga      120
cgctaaggaa tttgttctta cattggataa cactaatttc catgacactg ttaagaagca      180
cgatttcata gtcgttgaat tctacgcacc ttgggtgtgga cactgtaaga agctagcccc      240
agagtatgag aaggctgctt ctatcttgag cactcacgag ccaccagttg ttttggctaa      300
agttgatgcc aatgaggagc acaacaaaga cctcgcatcg gaaaatgatg ttaagggatt      360
ccaaccatt aagattttta ggaatggtgg aaagaacatt caagaatata aaggtccccg      420
tgaagctgaa ggtattgttg agtatattgaa aaaacaaagt ggccctgcat ccacagaaat      480
taaattctgct gatgatgcca ccgcttttct tggtgacaac aaagttgtta ttgtcggagt      540
tttccctaaa ttttctggtg aggagtacga taacttcatt gcattagcag agaagttgctg      600
ttctgactat gactttgctc acactttgaa tgccaaacac cttccaaagg gagactcatc      660
agtgtctggg cctgtgggta ggttatattaa gccatttgac gagctctttg ttgactcaaa      720
ggatttcaat gtagaagctc tagagaaatt cattgaagaa tccagtaccc caattgtgac      780
tgtcttcaac aatgagccta gcaatcacc ctttgttgct aaattcttta actctcccaa      840
cgcaaaggct atgttggtca tcaactttac taccgaagggt gctgaatctt tcaaaacaaa      900
ataccatgaa gtggctgagc aatacaaaca acagggaggt agctttcttg ttggagatgt      960
tgagtctagt caaggtgcct tccagtattt tggactgaag gaagaacaag tacctctaata      1020
tattattcag cataatgatg gcaagaagtt tttcaaacc aatttggaac ttgatcaact      1080
ccaacttggt ttgaaggcat acaaggatgg caaggttgaa ccatttgtca agtctgaacc      1140
tattcctgaa actaacaacg agcctgttaa agtgggtggt gggcaaactc ttgaggacgt      1200
tgttttcaag tctgggaaga atgttttgat agagttttat gtccttgggt gtggtcactg      1260
caagcagttg gctccaatct tggatgaagt tgctgtctca ttccaaagcg atgctgatgt      1320
tgttattgca aaactggatg caactgcaa cgatatccca accgacacct ttgatgtcca      1380
aggctatcca accttgact tcaggtcagc aagtggaaaa ctatcacaat acgacgggtg      1440
taggacaaag gaagacatca tagaattcat tgaaaagaac aaggataaaa ctggtgctgc      1500
tcatcaagaa gtagaacaac caaaagctgc tgctcagcca gaagcagaac aaccaaaga      1560

```

tgagctttga aaagttccgc ttggaggata tcggcacaca gtcactctgcg ggctttacaa 1620
 ctcttttgta tctcagaatc agaagttagg aaatcttagt gccaatctat ctatttttgc 1680
 gtttcatttt atctttttgg tttactctaa tgtattactg aataatgtga gttttggcgg 1740
 agtttagtac tggaactttt gtttctgtaa aaaaaaaaaa a 1781

<210> 58

<211> 1457

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự plasmid PromPlasto-PDISP-Plasto 3(gốc)UTR tổng hợp

<400> 58

ctggtatatt tatatgttgt caaataactc aaaaaccata aaagtttaag ttagcaagtg 60
 tgtacatttt tacttgaaca aaaatattca cctactactg ttataaatca ttattaaaca 120
 ttagagtaaa gaaatatgga tgataagaac aagagtagtg atattttgac aacaattttg 180
 ttgcaacatt tgagaaaatt ttgttgttct ctcttttcat tgggtcaaaaa caatagagag 240
 agaaaaagga agagggagaa taaaaacata atgtgagtat gagagagaaa gttgtacaaa 300
 agttgtacca aaatagttgt acaaatatca ttgaggaatt tgacaaaagc tacacaaata 360
 aggggttaatt gctgtaaata aataaggatg acgcattaga gagatgtacc attagagaat 420
 ttttggcaag tcattaataaa gaaagaataa attattttta aaattaaaag ttgagtcatt 480
 tgattaaaca tgtgattatt taatgaattg atgaaagagt tggattaaag ttgtattagt 540
 aattagaatt tgggtgtcaa ttttaatttg catttgatct tttcctatat attgccccat 600
 agagtcagtt aactcatttt tatatttcat agatcaaata agagaaataa cgggtatatta 660
 atccctccaa aaaaaaaaaa cgggtatatt actaaaaaat ctaagccacg taggaggata 720
 acaggatccc cgtaggagga taacatccaa tccaaccaat cacaacaatc ctgatgagat 780
 aaccacttt aagcccacgc atctgtggca catctacatt atctaaatca cacattcttc 840
 cacacatctg agccacacaa aaaccaatcc acatctttat caccattct ataaaaaatc 900
 acactttgtg agtctacact ttgattccct tcaaacacat acaaagagaa gagactaatt 960
 aattaattaa tcatcttgag agaaaatggc gaaaaacggt gcgattttcg gcttattggt 1020
 ttctcttctt gtgttggttc cttctcagat ctgagctcta agttaaaatg cttcttcgtc 1080
 tcctatttat aatatggttt gttattgtta attttgttct tgtagaagag ctttaattaat 1140
 cgttgttggt atgaaatact atttgtatga gatgaactgg tgtaatgtaa ttcatttaca 1200
 taagtggagt cagaatcaga atgtttcctc cataactaac tagacatgaa gacctgccgc 1260
 gtacaattgt cttatatttg aacaactaaa attgaacatc ttttgccaca actttataag 1320

tggttaatat agctcaaata tatggtcaag ttcaatagat taataatgga aatatcagtt 1380
 atcgaaattc attaacaatc aacttaacgt tattaactac taattttata tcatccctt 1440
 tgataaatga tagtaca 1457

<210> 59
 <211> 3206
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Catxet biểu hiện 540 từ HindIII tổng hợp

<400> 59
 aagcttgcta gcggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta 60
 tatattatg ttgtcaaata actcaaaaac cataaaagtt taagttagca agtgtgtaca 120
 tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag 180
 taaagaaata tggatgataa gaacaagagt agtgatattt tgacaacaat tttgttgcaa 240
 catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa 300
 aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt 360
 accaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataaggggt 420
 aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg 480
 caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta 540
 aacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag 600
 aatttggtgt caaatttaat ttgacatttg atcttttcct atatattgcc ccatagagtc 660
 agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaatccct 720
 ccaaaaaaaaa aaaacggtat atttactaaa aaatctaagc cacgtaggag gataacagga 780
 tccccgtagg aggataacat ccaatccaac caatcacaac aatcctgatg agataacca 840
 ctttaagccc acgcatctgt ggcacatcta cattatctaa atcacacatt ctccacaca 900
 tctgagccac acaaaaacca atccacatct ttatcaccca ttctataaaa aatcacactt 960
 tgtgagtcta cactttgatt cccttcaaac acatacaaag agaagagact aattaattaa 1020
 ttaatcatct tgagagaaaa tggcgaaaaa cgttgcgatt ttcggcttat tgttttctct 1080
 tcttggtgtg gttccttctc agatcttcgc tgacacaata tgtataggct accatgccaa 1140
 caactcaacc gacactgttg acacagtact tgagaagaat gtgacagtga cacactctgt 1200
 caacctactt gaggacagtc acaatggaaa actatgtcta ctaaaaggaa tagccccact 1260
 acaattgggt aattgcagcg ttgccggatg gatcttagga aaccagaat gcgaattact 1320
 gatttccaag gaatcatggt cctacattgt agaaacacca aatcctgaga atggaacatg 1380

ttacccaggg	tatttcgccg	actatgagga	actgagggag	caattgagtt	cagtatcttc	1440
atltgagaga	ttcgaaatat	tccccaaaga	aagctcatgg	cccaaccaca	ccgtaaccgg	1500
agtatcagca	tcatgctccc	ataatgggaa	aagcagtttt	tacagaaatt	tgctatggct	1560
gacggggaag	aatggtttgt	acccaaacct	gagcaagtcc	tatgtaaaca	acaaagagaa	1620
agaagtcctt	gtactatggg	gtgttcatca	ccgcctaac	ataggaacc	aaagggcact	1680
ctatcataca	gaaaatgctt	atgtctctgt	agtgtcttca	cattatagca	gaagattcac	1740
cccagaaata	gccaaaagac	caaagtaag	agatcaggaa	ggaagaatca	actactactg	1800
gactctgctg	gaacctgggg	atacaataat	atltgaggca	aatggaaatc	taatagcgcc	1860
atggtatgct	tttgactga	gtagaggctt	tggtatcagga	atcatcacct	caaatgcacc	1920
aatggatgaa	tgtgatgcga	agtgtcaaac	acctcaggga	gctataaaca	gcagtcttcc	1980
tttccagaat	gtacaccag	tcacaatagg	agagtgtcca	aagtatgtca	ggagtgcaaa	2040
attaaggatg	gttacaggac	taaggaacat	cccatccatt	caatccagag	gtttgtttgg	2100
agccattgcc	ggtttcattg	aaggggggtg	gactggaatg	gtagatgggt	ggatatggta	2160
tcatcatcag	aatgagcaag	gatctggcta	tgctgcagat	caaaaaagta	cacaaaatgc	2220
cattaacggg	attacaaaca	aggtcaattc	tgtaattgag	aaaatgaaca	ctcaattcac	2280
agctgtgggc	aaagagttca	acaaattgga	aagaaggatg	gaaaacttaa	ataaaaaagt	2340
tgatgatggg	tttctagaca	tttggacata	taatgcagaa	ttgttggttc	tactggaaaa	2400
tgaaaggact	ttggatttcc	atgactccaa	tgtgaagaat	ctgtatgaga	aagtaaaaag	2460
ccaattaaag	aataatgcca	aagaaatagg	aaacgggtgt	tttgagttct	atcacaagtg	2520
taacaatgaa	tgcatggaga	gtgtgaaaaa	tggtacctat	gactatccaa	aatattccga	2580
agaatcaaag	ttaaacaggg	agaaaattga	tgagtgaaa	ttggaatcaa	tgggagtata	2640
ccagattctg	gcatctact	caactgtcgc	cagttccctg	gttcttttgg	tctccctggg	2700
ggcaatcagc	ttctggatgt	gttccaatgg	gtctttgcag	tgtagaatat	gcatctaaga	2760
gctctaagtt	aaaatgcttc	ttcgtctcct	atttataata	tggtttgtta	ttgttaattt	2820
tgttcttgta	gaagagctta	attaatcggt	gttgttatga	aatactattt	gtatgagatg	2880
aactggtgta	atgtaattca	tttacataag	tgagtcaga	atcagaatgt	ttcctccata	2940
actaactaga	catgaagacc	tgccgcgtac	aattgtctta	tatttgaaca	actaaaattg	3000
aacatctttt	gccacaactt	tataagtgg	taatatagct	caaatatatg	gtcaagttca	3060
atagattaat	aatggaaata	tcagttatcg	aaattcatta	acaatcaact	taacgttatt	3120
aactactaat	tttatatcat	cccctttgat	aatgatagt	acaccaatta	ggaaggagca	3180

tgctcgaggc ctggctggcc gaattc

3206

<210> 60

<211> 1788

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn DraIII-Plasto(-84+1)-H1

A/Brisbane/59/07-SacI tổng hợp.

<400> 60

cactttgtga gtctacactt tgattccctt caaacacata caaagagaag agactaatta	60
attaattaat catcttgaga gaaaatgaaa gtaaaactac tggctcctgtt atgcacattt	120
acagctacat atgcagacac aatatgtata ggctaccatg ctaacaactc gaccgacact	180
gttgacacag tacttgaaaa gaatgtgaca gtgacacact ctgtcaacct gcttgagaac	240
agtcacaatg gaaaactatg tctattaaaa ggaatagccc cactacaatt gggtaattgc	300
agcgttgccg ggtggatctt aggaaaccca gaatgcgaat tactgatttc caaggagtca	360
tggctctaca ttgtagaaaa accaaatcct gagaatggaa catgttacct agggcatttc	420
gctgactatg aggaactgag ggagcaattg agttcagtat cttcatttga gaggttcgaa	480
atattcccca aagaaagctc atggcccaac cacaccgtaa ccggagtgtc agcatcatgc	540
tcccataatg gggaaagcag tttttacaga aatttgctat ggctgacggg gaagaatgg	600
ttgtacccaa acctgagcaa gtcctatgca aacaacaaag aaaaagaagt ctttgtacta	660
tgggggtgttc atcacccgcc aaacataggt gaccaaagg ccctctatca tacagaaaat	720
gcttatgtct ctgtagtgtc ttcacattat agcagaaaat tcaccccaga aatagccaaa	780
agacccaaag taagagatca agaaggaaga atcaattact actggactct gcttgaaccc	840
ggggatacaa taatatttga ggcaaagga aatctaatacg cccaagata tgctttcgca	900
ctgagtagag gctttggatc aggaatcatc aactcaaag caccatgga taaatgtgat	960
gcgaagtgcc aaacacctca gggagctata aacagcagtc ttcctttcca gaacgtacac	1020
ccagtcacaa taggagagtg tccaaagtat gtcaggagtg caaaattaag gatggttaca	1080
ggactaagga acatcccatc cattcaatcc agaggtttgt ttggagccat tgccggtttc	1140
attgaagggg ggtggactgg aatggtagat ggttggtatg gttatcatca tcagaatgag	1200
caaggatctg gctatgctgc agatcaaaaa agcacacaaa atgccattaa tgggattaca	1260
aacaaggtca attctgtaat tgagaaaatg aacactcaat tcacagcagt gggcaaagag	1320
ttcaacaaat tggaaagaag gatggaaaac ttgaataaaa aagttgatga tgggtttata	1380
gacatttgga catataatgc agaactgttg gttctactgg aaaatgaaag gactttggat	1440
ttccatgact ccaatgtgaa gaatctgtat gagaaagtaa aaagccagtt aaagaataat	1500

gctaaagaaa taggaaatgg gtgttttgag ttctatcaca agtgtaacga tgaatgcatg 1560
gagagtgtaa agaatggaac ttatgactat ccaaaatatt ccgaagaatc aaagttaaac 1620
agggagaaaa ttgatggagt gaaattggaa tcaatgggag tctatcagat tctggcgatc 1680
tactcaacag tcgccagttc tctggttctt ttgggtctccc tgggggcaat cagcttctgg 1740
atgtgttcca atgggtcttt acagtgtaga atatgcatct aagagctc 1788

<210> 61

<211> 3185

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Catxet biểu hiện 774 từ HindIII tổng hợp

<400> 61

aagcttgcta gcggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta 60
tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagttagca agtgtgtaca 120
tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag 180
taaagaaata tggatgataa gaacaagagt agtgatatct tgacaacaat tttgttgcaa 240
catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa 300
aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt 360
acaaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataaggggt 420
aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg 480
caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta 540
aacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag 600
aatttggtgt caaatttaat ttgacatttg atcttttcct atatattgcc ccatagagtc 660
agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaatccct 720
ccaaaaaaa aaaacggtat atttactaaa aaatctaagc cacgtaggag gataacagga 780
tccccgtagg aggataacat ccaatccaac caatcacaac aatcctgatg agataaccca 840
ctttaagccc acgcatctgt ggcacatcta cattatctaa atcacacatt cttccacaca 900
tctgagccac acaaaaacca atccacatct ttatcaccca ttctataaaa aatcacactt 960
tgtgagtcta cactttgatt cccttcaaac acatacaaag agaagagact aattaattaa 1020
ttaatcatct tgagagaaaa tgaaagtaaa actactgggtc ctgttatgca catttacagc 1080
tacatatgca gacacaatat gtataggcta ccatgctaac aactcgaccg aactgttga 1140
cacagtactt gaaaagaatg tgacagtgac aactctgtc aacctgcttg agaacagtca 1200
caatggaaaa ctatgtctat taaaaggaat agccccacta caattgggta attgcagcgt 1260

tgccgggtgg	atcttaggaa	accagaatg	cgaattactg	atttccaagg	agtcatggtc	1320
ctacattgta	gaaaaaccaa	atcctgagaa	tggaacatgt	taccaggggc	atttcgctga	1380
ctatgaggaa	ctgagggagc	aattgagttc	agtatcttca	tttgagaggt	tcgaaatatt	1440
cccaaagaa	agctcatggc	ccaaccacac	cgtaaccgga	gtgtcagcat	catgctccca	1500
taatggggaa	agcagttttt	acagaaattt	gctatggctg	acggggaaga	atggtttgta	1560
cccaaacctg	agcaagtcct	atgcaaacaa	caaagaaaaa	gaagtccttg	tactatgggg	1620
tgttcatcac	ccgccaacaa	taggtgacca	aaaggccctc	tatcatacag	aaaatgctta	1680
tgtctctgta	gtgtcttcac	attatagcag	aaaattcacc	ccagaaatag	ccaaaagacc	1740
caaagtaaga	gatcaagaag	gaagaatcaa	ttactactgg	actctgcttg	aaccggggga	1800
tacaataata	tttgaggcaa	atggaaatct	aatagcgcca	agatatgctt	tcgcactgag	1860
tagaggcttt	ggatcaggaa	tcatcaactc	aatgcacca	atggataaat	gtgatgcgaa	1920
gtgccaacaa	cctcagggag	ctataaacag	cagtcttcct	ttccagaacg	tacacccagt	1980
cacaatagga	gagtgtccaa	agtatgtcag	gagtgcaaaa	ttaaggatgg	ttacaggact	2040
aaggaacatc	ccatccattc	aatccagagg	tttgtttgga	gccattgccg	gtttcattga	2100
aggggggtgg	actggaatgg	tagatggttg	gtatggttat	catcatcaga	atgagcaagg	2160
atctggctat	gctgcagatc	aaaaaagcac	acaaaatgcc	attaatggga	ttacaaacaa	2220
ggtcaattct	gtaattgaga	aatgaacac	tcaattcaca	gcagtgggca	aagagttcaa	2280
caaattggaa	agaaggatgg	aaaacttgaa	taaaaaagtt	gatgatgggt	ttatagacat	2340
ttggacatat	aatgcagaac	tgttggttct	actggaaaat	gaaaggactt	tggatttcca	2400
tgactccaat	gtgaagaatc	tgtatgagaa	agtaaaaagc	cagttaaaga	ataatgctaa	2460
agaaatagga	aatgggtggt	ttgagttcta	tcacaagtgt	aacgatgaat	gcatggagag	2520
tgtaaagaat	ggaacttatg	actatccaaa	atattccgaa	gaatcaaagt	taaacaggga	2580
gaaaattgat	ggagtgaaat	tggaatcaat	gggagtctat	cagattctgg	cgatctactc	2640
aacagtcgcc	agttctctgg	ttcttttggt	ctccctgggg	gcaatcagct	tctggatgtg	2700
ttccaatggg	tctttacagt	gtagaatatg	catctaagag	ctctaagtta	aatgcttct	2760
tcgtctccta	tttataatat	ggtttggtat	tgtaaatttt	gttcttgtag	aagagcttaa	2820
ttaatcgttg	ttgttatgaa	atactatttg	tatgagatga	actggtgtaa	tgtaattcat	2880
ttacataagt	ggagtcagaa	tcagaatggt	tcctccataa	ctaactagac	atgaagacct	2940
gccgcgtaca	attgtcttat	atttgaacaa	ctaaaattga	acatcttttg	ccacaacttt	3000
ataagtgggt	aatatagctc	aaatatatgg	tcaagttcaa	tagattaata	atggaaatat	3060

cagttatcga aattcattaa caatcaactt aacgttatta actactaatt ttatatcatc 3120
 ccctttgata aatgatagta caccaattag gaaggagcat gctcgaggcc tggctggccg 3180
 aattc 3185
 <210> 62
 <211> 2065
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> Catxet biểu hiện số 828 tổng hợp, từ PacI đến AscI
 <400> 62
 ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct 60
 gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg 120
 ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tgggtccaaa gatggacccc 180
 caccacgag gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg 240
 attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag 300
 acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt 360
 tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcatc 420
 tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga 480
 tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg 540
 tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggt 600
 cttgggaaaa gaaagcttgc tggaggctgc tggtcagccc catacattac ttgttacgat 660
 tctgctgact ttcggcggtt gcaatatctc tacttctgct tgacgaggta ttgttgacctg 720
 tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa 780
 acagagtttt cccgtgggtt tcgaacttgg agaaagattg ttaagcttct gtatattctg 840
 cccaaatttg tcgggcccat ggttttcaca cctcagatac ttggacttat gcttttttgg 900
 atttcagcct ccagaggtga tattgtgcta actcagtctc cagccaccct gtctgtgact 960
 ccaggagata gtgtcagtct ttctgcagg gccagccaaa gtattagcaa caacctacac 1020
 tggtttcaac aaaaatcgca tgagtctcca aggcttctca tcaagtatgc ttcccagtcc 1080
 atatctggga tcccctccag gttcagtggc agtggatctg ggacagattt cactctcagt 1140
 atcaacagtg tgaagactga agattttgga atgtttttct gtcaacagag taacagctgg 1200
 cctctcacgt tcggtgatgg gacaaagctg gagctgaaac gggctgatgc tgcaccaact 1260
 gtatccatct tcccaccatc cagtgagcag ttaacatctg gaggtgcctc agtcgtgtgc 1320
 ttcttgaaca acttctaccc caaagacatc aatgtcaagt ggaagattga tggcagtgaa 1380

cgacaaaatg gcgtcctgaa cagttggact gatcaggaca gcaaagacag cacctacagc 1440
atgagcagca ccctcacggt gaccaaggac gagtatgaac gacataacag ctatacctgt 1500
gaggccactc acaagacatc aacttcaccc attgtcaaga gcttcaacag gaatgagtgt 1560
tagaggccta ttttcttttag tttgaattta ctgttattcg gtgtgcattt ctatgttttg 1620
tgagcggttt tctgtgctca gagtgtgttt attttatgta atttaatttc tttgtgagct 1680
cctgttttagc aggtcgtccc ttcagcaagg acacaaaaag attttaattt tattaaaaaa 1740
aaaaaaaaaa aagaccggga attcgatatac aagcttatcg acctgcagat cgttcaaaca 1800
tttggaataa aagtttctta agattgaatc ctgttgccgg tcttgcatg attatcatat 1860
aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgcatg acgttattta 1920
tgagatgggt ttttatgatt agagtccgc aattatacat ttaatacgcg atagaaaaca 1980
aaatatagcg cgcaactag gataaattat cgcgcgcgggt gtcactatg ttactagatt 2040
ctagagtctc aagcttcggc gcgcc 2065

<210> 63

<211> 3194

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế số 690 tổng hợp, từ HindIII đến EcoRI

<400> 63

aagcttgcta gcggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta 60
tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagtttagca agtgtgtaca 120
tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag 180
taaagaaata tggatgataa gaacaagagt agtgatatat tgacaacaat tttgttgcaa 240
catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa 300
aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt 360
acaaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataagggtt 420
aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttg 480
caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta 540
aacatgtgat tatttaataa attgatgaaa gagttggatt aaagttgtat tagtaattag 600
aatttggtgt caaatttaat ttgacatttg atcttttcct atatattgcc ccatagagtc 660
agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaattcct 720
ccaaaaaaa aaaacggtat atttactaaa aaatctaagc cacgtaggag gataacagga 780
tccccgtagg aggataacat ccaatccaac caatcacaac aatcctgatg agataacca 840

ctttaagccc	acgcatctgt	ggcacatcta	cattatctaa	atcacacatt	cttccacaca	900
tctgagccac	acaaaaacca	atccacatct	ttatcaccca	ttctataaaa	aatcacactt	960
tgtgagtcta	cacttttgatt	cccttcaaac	acatacaaag	agaagagact	aattaattaa	1020
ttaatcatct	tgagagaaaa	tggagaaaat	agtgtcttctt	cttgcaatag	tcagtcttgt	1080
taaaagtgat	cagatttgca	ttggttacca	tgcaaacaat	tcaacagagc	aggttgacac	1140
aatcatggaa	aagaacgtta	ctgttacaca	tgcccaagac	atactggaaa	agacacacaa	1200
cgggaagctc	tgcgatctag	atggagttaa	gcctctaatt	ttaagagatt	gtagtgtagc	1260
tggatggctc	ctcggaacc	caatgtgtga	cgaattcatc	aatgtaccgg	aatgggtctta	1320
catagtggag	aaggccaatc	caaccaatga	cctctgttac	ccagggcatt	tcgctgacta	1380
tgaggaactg	aggagcaat	tgagttcagt	atcttcattt	gagaggttcg	aaatattccc	1440
caaagaaagc	tcatggcca	accacaccgt	aaccggagtg	tcagcatcat	gctcccataa	1500
tggggaaagc	agtttttaca	gaaatttgct	atggctgacg	gggaagaatg	gtttgtaccc	1560
aaacctgagc	aagtcctatg	caaacaacaa	agaaaaagaa	gtccttgta	tatgggggtgt	1620
tcatcacccg	ccaaacatag	gtgacaaaa	ggccctctat	catacagaaa	atgcttatgt	1680
ctctgtagtg	tcttcacatt	atagcagaaa	attcacccca	gaaatagcca	aaagacccaa	1740
agtaagagat	caagaaggaa	gaatcaatta	ctactggact	ctgcttgaac	ccgggggatac	1800
aataatattt	gaggcaaatg	gaaatcta	agcgccaaga	tatgctttcg	cactgagtag	1860
aggctttgga	tcaggaatta	tgaaaagtga	attggaatat	ggtaactgca	acaccaagtg	1920
tcaaactcca	atgggggcca	taaactctag	tatgccattc	cacaacatac	accctctcac	1980
catcggggaa	tgccccaat	atgtgaaatc	aaacagatta	gtccttgcaa	cagggtcag	2040
aatagccct	caaagagaga	gcagaagaaa	aaagagagga	ctatttgag	ctatagcagg	2100
ttttatagag	ggaggatggc	agggaatggt	agatggttgg	tatgggtacc	accatagcaa	2160
tgagcagggg	agtgggtacg	ctgcagacaa	agaatccact	caaaaggcaa	tagatggagt	2220
caccaataag	gtcaactcaa	tcattgacaa	aatgaacact	cagtttgagg	ccgttggaag	2280
ggaattta	aacttagaaa	ggagaataga	gaatttaaac	aagaagatgg	aagacgggtt	2340
tctagatgtc	tggacttata	atgccgaact	tctggttctc	atggaaaatg	agagaactct	2400
agactttcat	gactcaaatg	ttaagaacct	ctacgacaag	gtccgactac	agcttaggga	2460
taatgcaaag	gagctgggta	acggttggtt	cgagttctat	cacaaatgtg	ataatgaatg	2520
tatggaaagt	ataagaaacg	gaacgtacaa	ctatccgcag	tattcagaag	aagcaagatt	2580
aaaaagagag	gaaataagtg	gggtaaaatt	ggaatcaata	ggaacttacc	aaatactgtc	2640
aattttattca	acagtggcca	gttccctagc	actggcaatc	atgatggctg	gtctatcttt	2700

atgggatgtgc tccaatggat cggtacaatg cagaatttgc atttaagagc tctaagttaa 2760
aatgcttctt cgtctcctat ttataatatg gtttgttatt gttaattttg ttcttgtaga 2820
agagcttaat taatcgttgt tggtatgaaa tactatattgt atgagatgaa ctggtgtaat 2880
gtaattcatt tacataagtg gagtcagaat cagaatgttt cctccataac taactagaca 2940
tgaagacctg ccgcgtacaa ttgtcttata tttgaacaac taaaattgaa catcttttgc 3000
cacaacttta taagtgggta atatagctca aatatatggg caagttcaat agattaataa 3060
tggaatatatc agttatcgaa attcattaac aatcaactta acgttattaa ctactaattt 3120
tatatcatcc cctttgataa atgatagtag accaattagg aaggagcatg ctcgaggcct 3180
ggctggccga attc 3194

<210> 64

<211> 3194

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế số 691 tổng hợp, từ HindIII đến EcoRI

<400> 64

aagcttgcta gcggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta 60
tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagtttagca agtgtgtaca 120
tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag 180
taaagaaata tggatgataa gaacaagagt agtgatattt tgacaacaat tttgttgcaa 240
catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa 300
aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt 360
acaaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataaggggt 420
aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg 480
caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta 540
aacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag 600
aatttggtgt caaatttaat ttgacatttg atcttttcct atatattgcc ccatagagtc 660
agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaatccct 720
ccaaaaaaa aaaacggtat atttactaaa aaatctaagc cacgtaggag gataacagga 780
tccccgtagg aggataacat ccaatccaac caatcacaac aatcctgatg agataacca 840
ctttaagccc acgcatctgt ggcacatcta cattatctaa atcacacatt cttccacaca 900
tctgagccac acaaaaacca atccacatct ttatcaccca ttctataaaa aatcacactt 960
tgtgagtcta cactttgatt cccttcaaac acatacaaag agaagagact aattaattaa 1020

ttaatcatct	tgagagaaaa	tggagaaaat	agtgttctt	cttgcaatag	tcagtcttgt	1080
taaaagtgat	cagatttgca	ttggttacca	tgcaaacaat	tcaacagagc	aggttgacac	1140
aatcatggaa	aagaacgtta	ctgttacaca	tgcccaagac	atactggaaa	agacacacaa	1200
cgggaagctc	tgcctattaa	aaggaatagc	cccactacaa	ttgggtaatt	gcagcgttgc	1260
cgggtggatc	ttaggaaacc	cagaatgcga	attactgatt	tccaaggagt	catggtccta	1320
cattgtagaa	aaaccaaata	ctgagaatgg	aacatgttac	ccagggcatt	tcgctgacta	1380
tgaggaactg	aggagcaat	tgagttcagt	atcttcattt	gagaggttcg	aaatattccc	1440
caaagaaagc	tcattggcca	accacaccgt	aaccggagtg	tcagcatcat	gtccccataa	1500
tggggaaagc	agtttttaca	gaaatttgct	atggctgacg	gggaagaatg	gtttgtaccc	1560
aaacctgagc	aagtcctatg	caaacaacaa	agaaaaagaa	gtccttgtag	tatgggggtg	1620
tcattcaccg	ccaacatag	gtgacaaaaa	ggccctctat	catacagaaa	atgcttatgt	1680
ctctgtagtg	tcttcacatt	atagcagaaa	attcacccca	gaaatagcca	aaagacccaa	1740
agtaagagat	caagaaggaa	gaatcaatta	ctactggact	ctgcttgaac	ccgggggatac	1800
aataatattt	gaggcaaatg	gaaatctaata	agcgccaaga	tatgctttcg	caactgagtag	1860
aggcttttga	tcaggaatca	tcaactcaaa	tgaccaaatg	gataaatgca	acaccaagtg	1920
tcaaactcca	atgggggcca	taaactctag	tatgccattc	cacaacatac	accctctcac	1980
catcggggaa	tgccccaat	atgtgaaatc	aaacagatta	gtccttgcaa	caggggtcag	2040
aaatagccct	caaagagaga	gcagaagaaa	aaagagagga	ctatttggag	ctatagcagg	2100
ttttatagag	ggaggatggc	agggaatggt	agatggttgg	tatgggtacc	accatagcaa	2160
tgagcagggg	agtgggtacg	ctgcagacaa	agaatccact	caaaaggcaa	tagatggagt	2220
caccaataag	gtcaactcaa	tcattgacaa	aatgaacact	cagtttgagg	ccgttggaag	2280
ggaatttaata	aacttagaaa	ggagaataga	gaatttaaac	aagaagatgg	aagacggggt	2340
tctagatgtc	tggacttata	atgccgaact	tctggttctc	atggaaaatg	agagaactct	2400
agactttcat	gactcaaatg	ttaagaacct	ctacgacaag	gtccgactac	agcttaggga	2460
taatgcaaag	gagctgggta	acggttggtt	cgagttctat	cacaaatgtg	ataatgaatg	2520
tatggaaagt	ataagaaacg	gaacgtacaa	ctatccgcag	tattcagaag	aagcaagatt	2580
aaaaagagag	gaaataagtg	gggtaaaatt	ggaatcaata	ggaacttacc	aaatactgtc	2640
aatttattca	acagtggcca	gttccctagc	actggcaatc	atgatggctg	gtctatcttt	2700
atggatgtgc	tccaatggat	cgttacaatg	cagaatttgc	atttaagagc	tctaagttaa	2760
aatgcttctt	cgtctcctat	ttataatatg	gtttgttatt	gttaattttg	ttctttagta	2820

agagcttaat taatcgttgt tgttatgaaa tactatattgt atgagatgaa ctggtgtaat 2880
gtaattcatt tacataagtg gagtcagaat cagaatgttt cctccataac taactagaca 2940
tgaagacctg ccgcgtacaa ttgtcttata tttgaacaac taaaattgaa catcttttgc 3000
cacaacttta taagtgggta atatagctca aatatatggg caagttcaat agattaataa 3060
tggaatatatc agttatcgaa attcattaac aatcaactta acgttattaa ctactaattt 3120
tatatcatcc cctttgataa atgatagtag accaattagg aaggagcatg ctcgaggcct 3180
ggctggccga attc 3194

<210> 65

<211> 3206

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế số 696 tổng hợp, từ HindIII đến EcoRI

<400> 65

aagcttgcta gcggcctcaa tggccctgca ggtcgactct agaggtaccc cgggctggta 60
tatttatatg ttgtcaaata actcaaaaac cataaaagtt taagtttagca agtgtgtaca 120
tttttacttg aacaaaaata ttcacctact actgttataa atcattatta aacattagag 180
taaagaaata tggatgataa gaacaagagt agtgatatatt tgacaacaat tttgttgcaa 240
catttgagaa aattttgttg ttctctcttt tcattgggtca aaaacaatag agagagaaaa 300
aggaagaggg agaataaaaa cataatgtga gtatgagaga gaaagttgta caaaagttgt 360
acaaaaatag ttgtacaaat atcattgagg aatttgacaa aagctacaca aataaggggt 420
aattgctgta aataaataag gatgacgcat tagagagatg taccattaga gaatttttgg 480
caagtcatta aaaagaaaga ataaattatt tttaaaatta aaagttgagt catttgatta 540
aacatgtgat tatttaatga attgatgaaa gagttggatt aaagttgtat tagtaattag 600
aatttggtgt caaatttaat ttgacatttg atcttttcct atatattgcc ccatagagtc 660
agttaactca tttttatatt tcatagatca aataagagaa ataacggtat attaatccct 720
ccaaaaaaaa aaaacggtat atttactaaa aaatctaagc cacgtaggag gataacagga 780
tccccgtagg aggataacat ccaatccaac caatcacaac aatcctgatg agataaccca 840
ctttaagccc acgcatctgt ggcacatcta cattatctaa atcacacatt cttccacaca 900
tctgagccac acaaaaacca atccacatct ttatcaccca ttctataaaa aatcacactt 960
tgtgagtcta cactttgatt cccttcaaac acatacaaag agaagagact aattaattaa 1020
ttaatcatct tgagagaaaa tggcgaaaaa cgttgcgatt ttcggcttat tgttttctct 1080
tcttgtgttg gttccttctc agatcttcgc tgacacaata tgtataggct accatgccaa 1140

caactcaacc	gacactgttg	acacagtact	tgagaagaat	gtgacagtga	cacactctgt	1200
caacctactt	gaggacagtc	acaatggaaa	actatgtcta	ctaaaaggaa	tagccccact	1260
acaattgggt	aattgcagcg	ttgccggatg	gatcttagga	aaccagaat	gcgaattact	1320
gatttccaag	gaatcatggt	cctacattgt	agaaacacca	aatcctgaga	atggaacatg	1380
ttaccaggg	agtttcaacg	actatgaaga	actgaaacac	ctattgagca	gaataaacca	1440
ttttgagaaa	attcaaata	tccccaaaag	ttcttgggtcc	gatcatgaag	cctcatcagg	1500
agtttagctca	gcatgtccat	acctgggaag	tccctccttt	tttagaaatg	tggtatggct	1560
tatcaaaaag	aacagtacat	acccaacaat	aaagaaaagc	tacaataata	ccaaccaaga	1620
ggatcttttg	gtactgtggg	gaattcacca	tcctaataat	gcggcagagc	agacaaggct	1680
atatcaaaac	ccaaccacct	atatttccat	tgggacatca	acactaaacc	agagattggt	1740
acaaaaata	gctactagat	ccaaagtaaa	cgggcaaagt	ggaaggatgg	agttcttctg	1800
gacaatttta	aaacctaata	atgcaatcaa	cttcgagagt	aatggaaatt	tcattgctcc	1860
agaatatgca	tacaaaattg	tcaagaaagg	ggactcagca	atcatcacct	caaatgcacc	1920
aatggatgaa	tgtgatgcca	agtgtcaaac	acctcaggga	gctataaaca	gcagtcttcc	1980
tttccagaat	gtacaccag	tcacaatagg	agagtgtcca	aagtatgtca	ggagtgcaaa	2040
attaaggatg	gttacaggac	taaggaacat	cccatccatt	caatccagag	gtttgtttgg	2100
agccattgcc	ggtttcattg	aaggggggtg	gactggaatg	gtagatgggt	ggtatgggta	2160
tcatcatcag	aatgagcaag	gatctggcta	tgctgcagat	caaaaaagta	cacaaaatgc	2220
cattaacggg	attacaaaca	aggtaaatc	tgtaattgag	aaaatgaaca	ctcaattcac	2280
agctgtgggc	aaagagttca	acaaattgga	aagaaggatg	gaaaacttaa	ataaaaaagt	2340
tgatgatggg	tttctagaca	tttgacata	taatgcagaa	ttgttggttc	tactggaaaa	2400
tgaaaggact	ttggatttcc	atgactccaa	tgtgaagaat	ctgtatgaga	aagtaaaaag	2460
ccaattaaag	aataatgcca	aagaaatagg	aaacgggtgt	tttgagttct	atcacaagtg	2520
taacaatgaa	tgcatggaga	gtgtgaaaaa	tggtacctat	gactatccaa	aatattccga	2580
agaatcaaag	ttaaacaggg	agaaaattga	tggagtgaaa	ttggaatcaa	tgggagtata	2640
ccagattctg	gcgatctact	caactgtcgc	cagttccctg	gttcttttgg	tctccctggg	2700
ggcaatcagc	ttctggatgt	gttccaatgg	gtctttgcag	tgtagaatat	gcatctaaga	2760
gctctaagtt	aaaatgcttc	ttcgtctcct	atattataata	tggtttggtta	ttgttaattt	2820
tgttcttgta	gaagagctta	attaatcggt	gttggttatga	aatactattt	gtatgagatg	2880
aactggtgta	atgtaattca	tttacataag	tggagtcaga	atcagaatgt	ttcctccata	2940
actaactaga	catgaagacc	tgccgcgtac	aattgtctta	tatttgaaca	actaaaattg	3000

aacatctttt gccacaactt tataagtggg taatatagct caaatatatg gtcaagttca 3060
 atagattaat aatggaaata tcagttatcg aaattcatta acaatcaact taacggttatt 3120
 aactactaat tttatatcat cccctttgat aaatgatagt acaccaatta ggaaggagca 3180
 tgctcgaggc ctggctggcc gaattc 3206

<210> 66

<211> 3058

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế 732 tổng hợp, từ PacI đến AscI

<400> 66

ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct 60
 gtcactttat tgagaagata gtggaaaagg aaggtggctc ctacaaatgc catcattgcg 120
 ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tgggtccaaa gatggacccc 180
 caccacgag gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg 240
 attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag 300
 acccttcctc tatataagga agttcatttc atttgagag gtattaaaat cttaataggt 360
 tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcatc 420
 tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga 480
 tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg 540
 tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggt 600
 cttgggaaaa gaaagcttgc tggaggctgc tggtcagccc catacattac ttgttacgat 660
 tctgctgact ttcggcggtt gcaatatctc tacttctgct tgacgaggta ttgttgctg 720
 tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa 780
 acagagtttt cccgtgggtt tcgaacttgg agaaagattg ttaagcttct gtatattctg 840
 cccaaatttg tcgggcccat gaaagtaaaa ctactggtcc tgttatgcac atttacagct 900
 acatatgcag acacaatatg tataggctac catgctaaca actcgaccga cactgttgac 960
 acagtacttg aaaagaatgt gacagtgaca cactctgtca acctgcttga gaacagtcac 1020
 aatggaaaac tatgtctatt aaaaggaata gcccactac aattgggtaa ttgcagcgtt 1080
 gccgggtgga tcttaggaaa ccagaaatgc gaattactga tttccaagga gtcatggtcc 1140
 tacattgtag aaaaaccaa tcttgagaat ggaacatgtt acccagggca tttcgctgac 1200
 tatgaggaac tgaggagca attgagttca gtatcttcat ttgagagggt cgaaatatc 1260
 cccaaagaaa gctcatggcc caaccacacc gtaaccggag tgtcagcatc atgctcccat 1320

aatggggaaa gcagttttta cagaaatttg ctatggctga cggggaagaa tggtttgtac 1380
ccaaacctga gcaagtccta tgcaaacaac aaagaaaaag aagtccttgt actatggggt 1440
gttcatcacc cgccaaacat aggtgaccaa aaggccctct atcatacaga aaatgcttat 1500
gtctctgtag tgtcttcaca ttatagcaga aaattcaccc cagaaatagc caaaagaccc 1560
aaagtaagag atcaagaagg aagaatcaat tactactgga ctctgcttga acccggggat 1620
acaataatat ttgaggcaaa tggaaatcta atagcgccaa gatatgcttt cgcactgagt 1680
agaggctttg gatcaggaat catcaactca aatgcaccaa tggataaatg tgatgcgaag 1740
tgccaaacac ctcagggagc tataaacagc agtcttcctt tccagaacgt acaccagtc 1800
acaataggag agtgtccaaa gtatgtcagg agtgcaaaat taaggatggg tacaggacta 1860
aggaacatcc catccattca atccagaggt ttgtttggag ccattgccgg tttcattgaa 1920
gggggggtgga ctggaatggg agatggttg tatggttatc atcatcagaa tgagcaagga 1980
tctggctatg ctgcagatca aaaaagcaca caaatgccaa ttaatgggat tacaacaag 2040
gtcaattctg taattgagaa aatgaacact caattcacag cagtgggcaa agagttcaac 2100
aaattggaaa gaaggatgga aaacttgaat aaaaaagttg atgatgggtt tatagacatt 2160
tggacatata atgcagaact gttggttcta ctggaaaatg aaaggacttt ggatttccat 2220
gactccaatg tgaagaatct gtatgagaaa gtaaaaagcc agttaagaa taatgctaaa 2280
gaaataggaa atgggtgttt tgagttctat cacaagtgtg acgatgaatg catggagagt 2340
gtaaagaatg gaacttatga ctatccaaaa tattccgaag aatcaaagtt aaacagggag 2400
aaaattgatg gagtgaatt ggaatcaatg ggagtctatc agattctggc gatctactca 2460
acagtcgcca gttctctggg tcttttgggc tccctggggg caatcagctt ctggatgtgt 2520
tccaatgggt ctttacagtg tagaatatgc atctaaaggc ctattttctt tagtttgaat 2580
ttactgttat tcggtgtgca tttctatgtt tggtgagcgg ttttctgtgc tcagagtgtg 2640
tttattttat gtaatttaat ttctttgtga gctcctgttt agcaggctgt cccttcagca 2700
aggacacaaa aagattttta ttttattaaa aaaaaaaaaa aaaaagaccg ggaattcgat 2760
atcaagctta tcgacctgca gatcggtcaa acatttgga ataaagtttc ttaagattga 2820
atcctgttgc cggctcttgcg atgattatca tataatttct gttgaattac gttaagcatg 2880
taataattaa catgtaatgc atgacgttat ttatgagatg ggtttttatg attagagtcc 2940
cgcaattata catttaatac gcgatagaaa acaaaatata gcgcgcaaac taggataaat 3000
tatcgcgcgc ggtgtcatct atgttactag attctagagt ctcaagcttc ggcgcgcc 3058

<210> 67

<211> 1719

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế số 787 tổng hợp

<400> 67

```

atggcgaaaa acgttgcgat ttctggctta ttgttttctc ttcttgtgtt gggttccttct      60
cagatcttcg ctgacacaat atgtataggc taccatgcta acaactcgac cgacactgtt      120
gacacagtac ttgaaaagaa tgtgacagtg acacactctg tcaacctgct tgagaacagt      180
cacaatggaa aactatgtct attaaaagga atagccccac tacaattggg taattgcagc      240
gttgccgggt ggatcttagg aaaccagaa tgccaattac tgatttccaa ggagtcattg      300
tcctacattg tagaaaaacc aaatcctgag aatggaacat gttaccagg gcatttcgct      360
gactatgagg aactgaggga gcaattgagt tcagtatctt catttgagag gttcgaaata      420
ttcccaaag aaagctcatg gccaaccac accgtaaccg gagtgtcagc atcatgctcc      480
cataatgggg aaagcagttt ttacagaaat ttgctatggc tgacggggaa gaatggtttg      540
tacccaaacc tgagcaagtc ctatgcaaac acaaagaaa aagaagtcct tgtactatgg      600
ggtgttcatt acccgccaaa cataggtgac caaaggccc tctatcatal agaaaatgct      660
tatgtctctg tagtgtcttc acattatagc agaaaattca cccagaaat agccaaaaga      720
cccaaagtaa gagatcaaga aggaagaatc aattactact ggactctgct tgaacccggg      780
gatacaataa tatttgaggc aaatggaaat ctaatagcgc caagatatgc ttctgcactg      840
agtagaggct ttggatcagg aatcatcaac tcaaatgcac caatggataa atgtgatgag      900
aagtgccaaa cacctcaggg agctataaac agcagtcttc ctttccagaa cgtacacca      960
gtcacaatag gagagtgtcc aaagtatgtc aggagtgcaa aattaaggat gggttacagga     1020
ctaaggaaca tcccatccat tcaatccaga ggtttggttg gagccattgc cggtttcatt     1080
gaaggggggt ggactggaat ggtagatggt tggtatggtt atcatcatca gaatgagcaa     1140
ggatctggct atgctgcaga tcaaaaagc acacaaaatg ccattaatgg gattacaaac     1200
aaggtcaatt ctgtaattga gaaaatgaac actcaattca cagcagtggg caaagagttc     1260
aacaattgg aaagaaggat ggaaaacttg aataaaaaag ttgatgatgg gtttatagac     1320
atttgacat ataatgcaga actgttggtt ctactggaaa atgaaaggac tttggatttc     1380
catgactcca atgtgaagaa tctgtatgag aaagtaaaaa gccagttaaa gaataatgct     1440
aaagaaatag gaaatgggtg ttttgagttc tatcacaagt gtaacgatga atgcatggag     1500
agtgtaaaga atggaactta tgactatcca aaatattccg aagaatcaaa gttaaacagg     1560
gagaaaattg atggagtga attggaatca atgggagtct atcagattct ggcgatctac     1620

```

tcaacagtcg ccagttctct ggttcttttg gtctccctgg gggcaatcag cttctggatg 1680

tgttccaatg ggtctttaca gtgtagaata tgcattctaa 1719

<210> 68
 <211> 3079
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Sản phẩm thiết kế số 733 tổng hợp, từ PacI đến AscI

<400> 68

ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct 60

gtcactttat tgagaagata gtggaaaagg aaggtggctc ctacaaatgc catcattgctg 120

ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tggccccaaa gatggacccc 180

caccacagag gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg 240

attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag 300

acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt 360

tttgataaaa gcgaacgtgg ggaaaccga accaaacctt cttctaaact ctctctcatc 420

tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga 480

tcgtgcttcg gcaccagtac aacgttttct ttactgaag cgaaatcaaa gatctctttg 540

tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggt 600

cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat 660

tctgctgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgctg 720

tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa 780

acagagtttt cccgtgggtt tcgaacttgg agaaagattg ttaagcttct gtatattctg 840

cccaaatttg tcgggcccat ggcgaaaaac gttgcgattt tcggcttatt gttttctctt 900

cttgtgttgg ttcttcttca gatcttcgct gacacaatat gtataggcta ccatgctaac 960

aactcgaccg aactgttga cacagtactt gaaaagaatg tgacagtgc acactctgtc 1020

aacctgcttg agaacagtca caatggaaaa ctatgtctat taaaaggaat agccccacta 1080

caattgggta attgcagcgt tgccgggtgg atcttaggaa acccagaatg cgaattactg 1140

atttccaagg agtcatggtc ctacattgta gaaaaaccaa atcctgagaa tggaacatgt 1200

taccagggc atttcgctga ctatgaggaa ctgagggagc aattgagttc agtatcttca 1260

tttgagaggt tcgaaatatt ccccaaagaa agctcatggc ccaaccacac cgtaaccgga 1320

gtgtcagcat catgctccca taatggggaa agcagttttt acagaaattt gctatggctg 1380

acggggaaga atggtttgta cccaaacctg agcaagtcct atgcaaacia caaagaaaaa 1440

gaagtccttg tactatgggg tgttcatcac ccgccaaaca taggtgacca aaaggccctc 1500
 tatcatacag aaaatgctta tgtctctgta gtgtcttcac attatagcag aaaattcacc 1560
 ccagaaatag ccaaaagacc caaagtaaga gatcaagaag gaagaatcaa ttactactgg 1620
 actctgcttg aaccgggga tacaataata tttgaggcaa atggaaatct aatagcgcca 1680
 agatatgctt tcgcactgag tagaggcttt ggatcaggaa tcatcaactc aaatgcacca 1740
 atggataaat gtgatgcgaa gtgccaaaca cctcaggag ctataaacag cagtcttctc 1800
 ttccagaacg tacaccagc cacaatagga gagtgtccaa agtatgtcag gagtgcaaaa 1860
 ttaaggatgg ttacaggact aaggaacatc ccatccattc aatccagagg tttgtttgga 1920
 gccattgccg gtttcattga aggggggtgg actggaatgg tagatgggtg gtatggttat 1980
 catcatcaga atgagcaagg atctggctat gctgcagatc aaaaaagcac acaaaatgcc 2040
 attaatggga ttacaaacaa ggtcaattct gtaattgaga aaatgaacac tcaattcaca 2100
 gcagtgggca aagagttcaa caaattggaa agaaggatgg aaaacttgaa taaaaaagtt 2160
 gatgatgggt ttatagacat ttggacatat aatgcagaac tgttggttct actggaaaat 2220
 gaaaggactt tggatttcca tgactccaat gtgaagaatc tgtatgagaa agtaaaaagc 2280
 cagttaaaga ataatgctaa agaaatagga aatgggtgtt ttgagttcta tcacaagtgt 2340
 aacgatgaat gcatggagag tgtaaagaat ggaacttatg actatccaaa atattccgaa 2400
 gaatcaaagt taaacaggga gaaaattgat ggagtgaat tggaatcaat gggagtctat 2460
 cagattctgg cgatctactc aacagtcgcc agttctctgg ttcttttgggt ctccctgggg 2520
 gcaatcagct tctggatgtg ttccaatggg tctttacagt gtagaatatg catctaaagg 2580
 cctattttct ttagtttgaa tttactgtta ttcggtgtgc atttctatgt ttggtgagcg 2640
 gttttctgtg ctgagagtgt gtttatttta tgtaatttaa tttctttgtg agctcctgtt 2700
 tagcaggctg tcccttcagc aaggacacaa aaagatttta atttatttaa aaaaaaaaaa 2760
 aaaaaagacc gggaattcga tatcaagctt atcgacctgc agatcgttca aacatttggc 2820
 aataaagttt ctaagattg aatcctgttg ccggtcttgc gatgattatc atataatttc 2880
 tgttgaatta cgtaagcat gtaataatta acatgtaatg catgacgtta tttatgagat 2940
 gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa aacaaaatat 3000
 agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta gattctagag 3060
 tctcaagctt cggcgcgcc 3079

<210> 69

<211> 3067

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế 734 tổng hợp, từ PacI đến AscI

<400> 69

ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct	60
gtcactttat tgagaagata gtggaaaagg aaggtggctc ctacaaatgc catcattgcg	120
ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tgggtccaaa gatggacccc	180
caccacagag gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg	240
attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag	300
acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt	360
tttgataaaa gcgaacgtgg ggaaaccgga accaaacctt cttctaaact ctctctcatc	420
tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga	480
tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg	540
tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcgggtgtgt	600
cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat	660
tctgctgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgctg	720
tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa	780
acagagtttt cccgtggttt tcgaacttgg agaaagattg ttaagcttct gtatattctg	840
cccaaatttg tcgggcccat ggagaaaata gtgcttcttc ttgcaatagt cagtcttggt	900
aaaagtgatc agatttgcat tggttaccat gcaaacaatt caacagagca ggttgacaca	960
atcatggaaa agaacgttac tgttacacat gcccaagaca tactggaaaa gacacacaac	1020
gggaagctct gcgatctaga tggagtgaag cctctaattt taagagattg tagtgtagct	1080
ggatggctcc tcgggaacc aatgtgtgac gaattcatca atgtaccgga atggtcttac	1140
atagtggaga aggccaatcc aaccaatgac ctctgttacc cagggcattt cgctgactat	1200
gaggaactga gggagcaatt gagttcagta tcttcatttg agaggttcga aatattcccc	1260
aaagaaagct catggcccaa ccacaccgta accggagtgt cagcatcatg ctcccataat	1320
ggggaaagca gtttttacag aaatttgcta tggctgacgg ggaagaatgg tttgtacca	1380
aacctgagca agtcctatgc aaacaacaaa gaaaaagaag tccttggtact atggggtgtt	1440
catcaccgcg caaacatagg tgaccaaaag gccctctatc atacagaaaa tgcttatgtc	1500
tctgtagtgt cttcacatta tagcagaaaa ttcaccccag aaatagccaa aagacccaaa	1560
gtaagagatc aagaaggaag aatcaattac tactggactc tgcttgaacc cggggataca	1620
ataatatattg aggcaaattg aaatctaata gcgccaaagat atgctttcgc actgagtaga	1680
ggctttggat caggaattat gaaaagtga ttggaatatg gtaactgcaa caccaagtgt	1740

caaactccaa tgggggcat aaactctagt atgccattcc acaacataca ccctctcacc 1800
 atcggggaat gcccacaaata tgtgaaatca aacagattag tccttgcaac agggctcaga 1860
 aatagccctc aaagagagag cagaagaaaa aagagaggac tatttgagc tatagcaggt 1920
 tttatagagg gaggatggca gggaaatggta gatgggtgggt atgggtacca ccatagcaat 1980
 gagcagggga gtgggtacgc tgcagacaaa gaatccactc aaaaggcaat agatggagtc 2040
 accaataagg tcaactcaat cattgacaaa atgaacactc agtttgaggc cgttggaagg 2100
 gaatttaata acttagaaag gagaatagag aatttaaaca agaagatgga agacgggttt 2160
 ctagatgtct ggacttataa tgccgaactt ctggttctca tggaaaatga gagaactcta 2220
 gactttcatg actcaaagt taagaacctc tacgacaagg tccgactaca gcttagggat 2280
 aatgcaaagg agctgggtaa cggttggttc gagttctatc acaaagtga taatgaatgt 2340
 atggaaagta taagaaacgg aacgtacaac tatccgcagt attcagaaga agcaagatta 2400
 aaaagagagg aaataagtgg ggtaaaattg gaatcaatag gaacttacca aatactgtca 2460
 atttattcaa cagtggcgag ttccctagca ctggcaatca tgatggctgg tctatcttta 2520
 tggatgtgct ccaatggatc gttacaatgc agaatttgca tttaaaggcc tattttcttt 2580
 agtttgaatt tactgttatt cggtgtgcat ttctatgttt ggtgagcggg tttctgtgct 2640
 cagagtgtgt ttattttatg taatttaatt tctttgtgag ctctgttta gcaggctgctc 2700
 ccttcagcaa ggacacaaaa agattttaat tttattaaaa aaaaaaaaaa aaaagaccgg 2760
 gaattcgata tcaagcttat cgacctgcag atcgttcaaa catttggaac taaagtttct 2820
 taagattgaa tcctgttgcc ggtcttgcca tgattatcat ataatttctg ttgaattacg 2880
 ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg gtttttatga 2940
 ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag cgcgcaaact 3000
 aggataaatt atcgcgcgcg gtgtcatcta tgttactaga ttctagagtc tcaagcttcg 3060
 gcgcgcc 3067

<210> 70

<211> 1791

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn DraIII-Plasto(-84+1)-H3
A/Brisbane/10/07-SacI tổng hợp

<400> 70

cactttgtga gtctacactt tgattccctt caaacacata caaagagaag agactaatta 60
 attaattaat catcttgaga gaaaatgaag actatcattg ctttgagcta cattctatgt 120

ctggttttca ctcaaaaact tcccggaaat gacaacagca cggcaacgct gtgccttggg 180
 caccatgcag taccaaacgg aacgatagtg aaaacaatca cgaatgacca aattgaagtt 240
 actaatgcta ctgagctggg tcagagttcc tcaacagggtg aaatatgcga cagtcctcat 300
 cagatccttg atggagaaaa ctgcacacta atagatgctc tattgggaga ccctcagtgt 360
 gatggcttcc aaaataagaa atgggacctt tttgttgaac gcagcaaagc ctacagcaac 420
 tgttaccctt atgatgtgcc ggattatgcc tcccttaggt cactagttgc ctcatccggc 480
 aactggagt ttaacaatga aagtttcaat tggactggag tcaactcaaaa cggaacaagc 540
 tctgcttgca taaggagatc taataacagt ttcttttagta gattgaattg gttgaccac 600
 ttaaaattca aataccagc attgaacgtg actatgccaa acaatgaaaa atttgacaaa 660
 ttgtacattt ggggggttca ccacccgggt acggacaatg accaaatctt cctgtatgct 720
 caagcatcag gaagaatcac agtctctacc aaaagaagcc acaaaactgt aatcccgaat 780
 atcggatcta gaccagagt aaggaatatc cccagcagaa taagcatcta ttggacaata 840
 gtaaaaccgg gagacatact tttgattaac agcacaggga atctaattgc tcctaggggt 900
 tacttcaaaa tacgaagtgg gaaaagctca ataatgagat cagatgcacc cattggcaaaa 960
 tgcaattctg aatgcatcac tccaaacgga agcattccca atgacaaacc attccaaaat 1020
 gtaaacagga tcacatacgg ggcctgtccc agatatgtta agcaaaacac tctgaaattg 1080
 gcaacaggga tgcgaaatgt accagagaaa caaactagag gcatatttgg cgcaatcgcg 1140
 ggtttcatag aaaatgggtg ggagggaatg gtggatgggtt ggtatgggtt caggcatcaa 1200
 aattctgagg gaataggaca agcagcagat ctcaaaagca ctcaagcagc aatcgatcaa 1260
 atcaatggga agctgaatag gttgatcggg aaaaccaacg agaaattcca tcagattgaa 1320
 aaagagttct cagaagtcga agggagaatc caggaccttg agaaatatgt tgaggacacc 1380
 aaaatagatc tctggtcata caacgcggag cttcttggtg ccctggagaa ccaacataca 1440
 attgatctaa ctgactcaga aatgaacaaa ctgtttgaaa aaacaaagaa gcaactgagg 1500
 gaaaatgctg aggatatggg caatggttgt ttcaaaatat accacaaatg tgacaatgcc 1560
 tgcataggat caatcagaaa tggaacttat gaccacgatg tatacagaga tgaagcatta 1620
 aacaaccggg tccagatcaa gggcgttgag ctgaagtcag gatacaaaga ttggatacta 1680
 tggatttctt ttgcatatc atgttttttg ctttgtgttg ctttgttggg gttcatcatg 1740
 tgggcctgcc aaaaaggcaa cattaggtgc aacatttgca tttgagagct c 1791

<210> 71

<211> 3085

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế 736 tổng hợp, từ PacI đến AscI

<400> 71

ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct	60
gtcactttat tgagaagata gtggaaaagg aaggtggctc ctacaaatgc catcattgcg	120
ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tggccccaaa gatggacccc	180
caccacagag gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg	240
attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag	300
acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt	360
tttgataaaa gcgaacgtgg ggaaacccga accaaacctt cttctaaact ctctctcatc	420
tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga	480
tcgtgcttcg gcaccagtac aacgttttct ttcactgaag cgaaatcaaa gatctctttg	540
tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcgggtgtgt	600
cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat	660
tctgctgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgctg	720
tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa	780
acagagtttt cccgtggttt tcgaacttgg agaaagattg ttaagcttct gtatattctg	840
cccaaatttg tcgggcccat ggcgaaaaac gttgcgattt tcggcttatt gttttctctt	900
cttgtgttgg ttccttctca gatcttcgct caaaaacttc ccggaaatga caacagcacg	960
gcaacgctgt gccttgggca ccatgcagta ccaaacggaa cgatagtga aacaatcacg	1020
aatgaccaa ttgaagttac taatgctact gagctgggtc agagttcctc aacaggtgaa	1080
atatgcgaca gtcctcatca gatccttgat ggagaaaact gcacactaat agatgctcta	1140
ttgggagacc ctgagtgtga tggcttccaa aataagaaat gggacctttt tgttgaacgc	1200
agcaaagcct acagcaactg ttacccttat gatgtgccgg attatgcctc ccttaggtca	1260
ctagttgcct catccggcac actggagttt aacaatgaaa gtttcaattg gactggagtc	1320
actcaaaacg gaacaagctc tgcttgcata aggagatcta ataacagttt ctttagtaga	1380
ttgaatttgt tgaccactt aaaattcaaa taccagcat tgaacgtgac tatgccaaac	1440
aatgaaaaat ttgacaaatt gtacatttgg ggggttcacc acccgggtac ggacaatgac	1500
caaactttcc tgtatgctca agcatcagga agaatacacag tctctaccaa aagaagccaa	1560
caaactgtaa tcccgaatat cggatctaga cccagagtaa ggaatatccc cagcagaata	1620
agcatctatt ggacaatagt aaaaccggga gacatacttt tgattaacag cacagggaat	1680
ctaattgctc ctaggggtta cttcaaaata cgaagtggga aaagctcaat aatgagatca	1740

gatgcaccca ttggcaaatg caattctgaa tgcatcactc caaacggaag cattcccaat 1800
gacaaaccat tccaaaatgt aaacaggatc acatacgggg cctgtcccag atatgttaag 1860
caaaacactc tgaaattggc aacagggatg cgaaatgtac cagagaaaca aactagaggc 1920
atatttggcg caatcgcggg ttcatagaa aatgggtggg agggaatggt ggatgggttg 1980
tatggtttca ggcatacaaa ttctgaggga ataggacaag cagcagatct caaaagcact 2040
caagcagcaa tcgatcaaat caatgggaag ctgaataggt tgatcgggaa aaccaacgag 2100
aaattccatc agattgaaaa agagtcttca gaagtcgaag ggagaatcca ggaccttgag 2160
aaatatgttg aggacaccaa aatagatctc tggcataca acgcggagct tcttggtgcc 2220
ctggagaacc aacatacaat tgatctaact gactcagaaa tgaacaaact gtttgaaaaa 2280
acaaagaagc aactgaggga aaatgctgag gatatgggca atggttggtt caaaatatac 2340
cacaaatgtg acaatgcctg cataggatca atcagaaatg gaacttatga ccacgatgta 2400
tacagagatg aagcattaaa caaccggttc cagatcaagg gcgttgagct gaagtcagga 2460
taciaaagatt ggatactatg gatttccttt gccatatcat gttttttgct ttgtgttgct 2520
ttgttggggg tcatcatgtg ggctgcca aaaggcaaca ttaggtgcaa catttgcat 2580
tgaaggccta ttttctttag ttgaattta ctgttattcg gtgtgcattt ctatgtttgg 2640
tgagcgggtt tctgtgctca gagtgtgtt attttatgta atttaatttc tttgtgagct 2700
cctgttttagc aggtcgtccc ttcagcaagg acacaaaaag attttaattt tattaaaaaa 2760
aaaaaaaaaa aagaccggga attcgatatc aagcttatcg acctgcagat cgttcaaaca 2820
tttggaata aagtttctta agattgaatc ctgttgccgg tcttgcgatg attatcatat 2880
aatttctggt gaattacggt aagcatgtaa taattaacat gtaatgcatg acgttattta 2940
tgagatgggt ttttatgatt agagtccgc aattatacat ttaatacgcg atagaaaaca 3000
aaatatagcg cgcaactag gataaattat cgcgcgcggg gtcacatg ttactagatt 3060
ctagagtctc aagcttcggc gcgcc 3085

<210> 72

<211> 3088

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế 737 tổng hợp, từ PacI đến AscI

<400> 72

ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct 60
gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg 120
ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tgggtcccaa gatggacccc 180

caccacgag gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg 240
 attgatgtga tatctccact gacgtaaggg atgacgcaca atccactat ccttcgcaag 300
 acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaataggt 360
 tttgataaaa gcgaacgtgg ggaaaccgga accaaacctt cttctaaact ctctctcatc 420
 tctcttaaag caaacttctc tcttgtcttt cttgcgtgag cgatcttcaa cgttgtcaga 480
 tcgtgcttcg gcaccagtac aacgttttct ttactgaag cgaaatcaaa gatctctttg 540
 tggacacgta gtgcggcgcc attaaataac gtgtacttgt cctattcttg tcggtgtggt 600
 cttgggaaaa gaaagcttgc tggaggctgc tgttcagccc catacattac ttgttacgat 660
 tctgctgact ttcggcgggt gcaatatctc tacttctgct tgacgaggta ttgttgctg 720
 tacttctttc ttcttcttct tgctgattgg ttctataaga aatctagtat tttctttgaa 780
 acagagtttt cccgtgggtt tcgaacttgg agaaagattg ttaagcttct gtatattctg 840
 cccaaatttg tcgggcccat ggcgaaaaac gttgcgattt tcggcttatt gttttctctt 900
 cttgtgttgg ttcttctca gatcttcgct caaaaacttc ccggaaatga caacagcacg 960
 gcaacgctgt gccttgggca ccatgcagta ccaaacggaa cgatagtga aacaatcacg 1020
 aatgaccaa ttgaagttac taatgctact gagctggttc agagttcctc aacaggtgaa 1080
 atatgcgaca gtcctcatca gatccttgat ggagaaaact gcacactaat agatgctcta 1140
 ttgggagacc ctcaagtgtga tggcttccaa aataagaaat gggacctttt tgttgaacgc 1200
 agcaaagcct acagcaactg ttacccttat gatgtgccg attatgcctc ccttaggtca 1260
 ctagttgcct catccggcac actggagttt aacaatgaaa gtttcaattg gactggagtc 1320
 actcaaaacg gaacaagctc tgcttgcata aggagatcta ataacagttt ctttagtaga 1380
 ttgaattggt tgaccactt aaaattcaaa taccagcat tgaacgtgac tatgccaaac 1440
 aatgaaaaat ttgacaaatt gtacatttgg ggggttcacc acccggtac ggacaatgac 1500
 caaatcttcc tgtatgctca agcatcagga agaatacag tctctaccaa aagaagccaa 1560
 caaactgtaa tcccgaatat cggatctaga ccagagtaa ggaatatccc cagcagaata 1620
 agcatctatt ggacaatagt aaaaccgga gacatacttt tgattaacag cacagggaat 1680
 ctaattgctc ctaggggtta cttcaaaata cgaagtggga aaagctcaat aatgagatca 1740
 gatgcacca ttggcaaatg caattctgaa tgcactctc caaacggaag cattcccaat 1800
 gacaaacat tccaaaatgt aaacaggatc acatacgggg cctgtcccag atatgttaag 1860
 caaaacactc tgaaattggc aacagggatg cgaaatgtac cagagaaaca aactagaggc 1920
 atatttggcg caatcgcggg tttcatagaa aatggttggg agggaatggt ggatggttgg 1980

tatggttttca ggcatacaaaa ttctgaggga ataggacaag cagcagatct caaaagcact	2040
caagcagcaa tcgatcaaat caatgggaag ctgaataggt tgatcgggaa aaccaacgag	2100
aaattccatc agattgaaaa agagtctca gaagtcgaag ggagaatcca ggaccttgag	2160
aatatgttg aggacaccaa aatagatctc tggatcataca acgcggagct tcttggtgcc	2220
ctggagaacc aacatacaat tgatctaact gactcagaaa tgaacaaact gtttgaaaaa	2280
acaaagaagc aactgaggga aaatgctgag gatatgggca atggttggtt caaaatatac	2340
cacaaatgtg acaatgcctg cataggatca atcagaaatg gaacttatga ccacgatgta	2400
tacagagatg aagcattaaa caaccggttc cagatcaagg gcgttgagct gaagtcaata	2460
ggaacttacc aaatactgtc aattttattca acagtggcga gttccctagc actggcaatc	2520
atgatggctg gtctatcttt atggatgtgc tccaatggat cggtacaatg cagaatttgc	2580
atttaaaggc ctatcttctt tagtttgaat ttactgttat tcggtgtgca tttctatgtt	2640
tggtgagcgg ttttctgtgc tcagagtgtg tttatcttat gtaatttaat ttctttgtga	2700
gctcctgttt agcaggtcgt cccttcagca aggacacaaa aagattttaa ttttattaaa	2760
aaaaaaaaa aaaaagaccg ggaattcgat atcaagctta tcgacctgca gatcggtcaa	2820
acatttgcca ataaagtctt ttaagattga atcctgttgc cggtcttgcg atgattatca	2880
tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat	2940
ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa	3000
acaaaatata gcgcgcaaac taggataaat tatcgcgcgcg ggtgtcatct atgttactag	3060
attctagagt ctcaagcttc ggcgcgcc	3088

<210> 73

<211> 1845

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn DraIII-Plasto(-84+1)-HA B/Florida/4/06-SacI tổng hợp

<400> 73

cactttgtga gtctacactt tgattccctt caaacacata caaagagaag agactaatta	60
attaattaat catcttgaga gaaaatgaag gcaataattg tactactcat ggtagtaaca	120
tccaatgcag atcgaatctg cactggaata acatcttcaa actcacctca tgtgggtcaaa	180
acagccactc aaggggaggt caatgtgact ggtgtgatac cactaacaac aacaccaaca	240
aatctttatt ttgcaaactc caaaggaaca aggaccagag ggaaactatg cccagactgt	300
ctcaactgca cagatctgga tgtggctttg ggcagaccaa tgtgtgtggg gaccacacct	360
tcggcgaagg cttcaatact ccacgaagtc aaacctgtta catccgggtg ctttcctata	420

atgcacgaca gaacaaaaat caggcaacta cccaatcttc tcagaggata tgaaaatatc 480
 aggctatcaa cccaaaacgt catcgatgcg gaaaaggcac caggaggacc ctacagactt 540
 ggaacctcag gatcttgccc taacgctacc agtaagagcg gatttttcgc aacaatggct 600
 tgggctgtcc caaaggacaa caacaaaaat gcaacgaacc cactaacagt agaagtacca 660
 tacatttgta cagaagggga agaccaaadc actgtttggg gggtccattc agataacaaa 720
 acccaaatga agaacctcta tggagactca aatcctcaaa agttcacctc atctgctaata 780
 ggagtaacca cacactatgt ttctcagatt ggcagcttcc cagatcaaac agaagacgga 840
 ggactaccac aaagcggcag gattgttggt gattacatga tgcaaaaacc tgggaaaaca 900
 ggaacaattg tctaccaaag aggtgttttg ttgcctcaaa aggtgtggtg cgcgagtggc 960
 aggagcaaag taataaaagg gtccttgccct ttaattgggtg aagcagattg ccttcatgaa 1020
 aaatacgggtg gattaaacaa aagcaagcct tactacacag gagaacatgc aaaagccata 1080
 ggaaattgcc caatatgggt gaaaacacct ttgaagctcg ccaatggaac caaatataga 1140
 cctcctgcaa aactattaaa ggaaaggggt ttcttcggag ctattgctgg tttcctagaa 1200
 ggaggatggg aaggaatgat tgcaggctgg cacggatata catctcacgg agcacatgga 1260
 gtggcagtgg cggcggacct taagagtacg caagaagcta taaacaagat aacaaaaaat 1320
 ctcaattctt tgagtgagct agaagtaaag aatcttcaaa gactaagtgg tgccatggat 1380
 gaactccaca acgaaatact cgagctggat gagaaagtgg atgatctcag agctgacact 1440
 ataagctcgc aaatagaact tgcagtcttg ctttccaacg aaggaataat aaacagtgaa 1500
 gatgagcatc tattggcact tgagagaaaa ctaaagaaaa tgctgggtcc ctctgctgta 1560
 gagataggaa atggatgctt cgaaacaaaa cacaagtgca accagacctg cttagacagg 1620
 atagctgctg gcacctttta tgcaggagaa ttttctctcc ccacttttga ttcactgaac 1680
 attactgctg catcttttaa tgatgatgga ttggataacc atactatact gctctattac 1740
 tcaactgctg cttctagttt ggctgtaaca ttgatgctag ctatttttat tgtttatatg 1800
 gtctccagag acaacgtttc atgctccatc tgtctataag agctc 1845

<210> 74

<211> 3142

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế 739 tổng hợp, từ PacI đến AscI

<400> 74

ttaattaaga attcgagctc caccgcgga acctcctcgg attccattgc ccagctatct 60

gtcactttat tgagaagata gtggaaaagg aaggtggctc ctacaaatgc catcattgcg 120

ataaaggaaa	ggccatcggt	gaagatgcct	ctgccgacag	tggtcccaa	gatggacccc	180
caccacagag	gagcatcggt	gaaaaagaag	acgttccaac	cacgtcttca	aagcaagtgg	240
attgatgtga	tatctccact	gacgtaaggg	atgacgcaca	atcccactat	ccttcgcaag	300
acccttcctc	tatataagga	agttcatttc	atttgagag	gtattaaaat	cttaataggt	360
tttgataaaa	gcgaacgtgg	ggaaacccga	accaaaccct	cttctaaact	ctctctcatc	420
tctcttaaag	caaacttctc	tcttgctctt	cttgcgtag	cgatcttcaa	cgttgctcaga	480
tcgtgcttcg	gcaccagtac	aacgttttct	ttcactgaag	cgaaatcaaa	gatctctttg	540
tggacacgta	gtgcggcgcc	attaaataac	gtgtacttgt	cctattcttg	tcggtgtggt	600
cttgggaaaa	gaaagcttgc	tggaggctgc	tggtcagccc	catacattac	ttgttacgat	660
tctgctgact	ttcggcggt	gcaatatctc	tacttctgct	tgacgaggta	ttgttgcttg	720
tacttcttct	ttcttcttct	tgctgattgg	ttctataaga	aatctagtat	tttctttgaa	780
acagagtttt	cccgtggttt	tcgaacttgg	agaaagattg	ttaagcttct	gtatattctg	840
cccaaatttg	tcgggcccac	ggcgaaaaac	gttgcgattt	tcggcttatt	gttttctctt	900
cttggtgttg	ttccttctca	gatcttcgct	gatcgaatct	gcactggaat	aacatcttca	960
aactcacctc	atgtggtcaa	aacagccact	caaggggagg	tcaatgtgac	tggtgtgata	1020
ccactaacia	caacaccaac	aaaatcttat	tttgcaaata	tcaaaggaa	aaggaccaga	1080
gggaaactat	gccagactg	tctcaactgc	acagatctgg	atgtggcttt	gggcagacca	1140
atgtgtgtgg	ggaccacacc	ttcggcgaag	gcttcaatac	tccacgaagt	caaacctgtt	1200
acatccgggt	gctttcctat	aatgcacgac	agaacaaaaa	tcaggcaact	acccaatctt	1260
ctcagaggat	atgaaaatat	caggctatca	acccaaaacg	tcacgatgc	ggaaaaggca	1320
ccaggaggac	cctacagact	tggaacctca	ggatcttgcc	ctaacgctac	cagtaagagc	1380
ggatttttct	caacaatggc	ttgggctgtc	ccaaaggaca	acaacaaaaa	tgcaacgaac	1440
ccactaacag	tagaagtacc	atacatttgt	acagaagggg	aagaccaa	caactgtttg	1500
gggttccatt	cagataacia	aacccaaatg	aagaacctct	atggagactc	aaatcctcaa	1560
aagttcacct	catctgctaa	tggagtaacc	acacactatg	tttctcagat	tggcagcttc	1620
ccagatcaaa	cagaagacgg	aggactacca	caaagcggca	ggattgttgt	tgattacatg	1680
atgcaaaaaa	ctgggaaaaa	aggaacaatt	gtctacaaaa	gaggtgtttt	gttgcttcaa	1740
aaggtgtggt	gcgcgagtgg	caggagcaaa	gtaataaaa	ggctcttgcc	tttaattggt	1800
gaagcagatt	gccttcatga	aaaatacggg	ggattaaaca	aaagcaagcc	ttactacaca	1860
ggagaacatg	caaaagccat	aggaaattgc	ccaatatggg	tgaaaacacc	tttgaagctc	1920
gccaatggaa	ccaaatatag	acctcctgca	aaactattaa	aggaaagggg	tttcttcgga	1980

gctattgctg gtttcctaga aggaggatgg gaaggaatga ttgcaggctg gcacggatac 2040
 acatctcacg gagcacatgg agtggcagtg gcggcggacc ttaagagtac gcaagaagct 2100
 ataaacaaga taacaaaaaa tctcaattct ttgagtgagc tagaagtaaa gaatcttcaa 2160
 agactaagtg gtgccatgga tgaactccac aacgaaatac tcgagctgga tgagaaagtg 2220
 gatgatctca gagctgacac tataagctcg caaatagaac ttgcagtctt gctttccaac 2280
 gaaggaataa taaacagtga agatgagcat ctattggcac ttgagagaaa actaaagaaa 2340
 atgctgggtc cctctgctgt agagatagga aatggatgct tcgaaaccaa acacaagtgc 2400
 aaccagacct gcttagacag gatagctgct ggcaccttta atgcaggaga attttctctc 2460
 cccacttttg attcactgaa cattactgct gcatctttaa atgatgatgg attggataac 2520
 catactatac tgctctatta ctcaactgct gcttctagtt tggctgtaac attgatgcta 2580
 gctattttta ttgtttatat ggtctccaga gacaacgttt catgctccat ctgtctataa 2640
 aggctatatt tctttagttt gaatttactg ttattcgggtg tgcatttcta tgtttggtga 2700
 gcggttttct gtgctcagag tgtgtttatt ttatgtaatt taatttcttt gtgagctcct 2760
 gtttagcagg tcgtcccttc agcaaggaca caaaaagatt ttaattttat taaaaaaaaa 2820
 aaaaaaaaaag accgggaatt cgatatcaag cttatcgacc tgcagatcgt tcaaacattt 2880
 ggcaataaag tttcttaaga ttgaatcctg ttgccgggtc tgcatgatt atcatataat 2940
 ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg ttatttatga 3000
 gatgggtttt tatgattaga gtcccgcaat tatacattta atacgcgata gaaaacaaaa 3060
 tatagcgcg c aaactaggat aaattatcgc gcgcgggtgc atctatgtta ctagattcta 3120
 gagtctcaag cttcggcgcg cc 3142

<210> 75

<211> 3142

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế 745, từ PacI đến AscI

<400> 75

ttaattaaga attcgagctc caccgcggaa acctcctcgg attccattgc ccagctatct 60
 gtcactttat tgagaagata gtggaaaagg aagggtggctc ctacaaatgc catcattgcg 120
 ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tgggtccaaa gatggacccc 180
 caccacagag gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg 240
 attgatgtga tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag 300
 acccttcctc tatataagga agttcatttc atttggagag gtattaaaat cttaatagggt 360

tttgataaaa	gcgaacgtgg	ggaaacccga	accaaaccctt	cttctaaact	ctctctcatc	420
tctcttaaag	caaactttctc	tcttgtcttt	cttgcgtagag	cgatcttcaa	cgttgtcaga	480
tcgtgcttcg	gcaccagtac	aacgttttct	ttcactgaag	cgaaatcaaa	gatctctttg	540
tggacacgta	gtgcggcgcc	attaaataac	gtgtacttgt	cctattcttg	tcggtgtggt	600
cttgggaaaa	gaaagcttgc	tggaggctgc	tgttcagccc	catacattac	ttgttacgat	660
tctgctgact	ttcggcggggt	gcaatatctc	tactttctgct	tgacgaggta	ttgttgacctg	720
tactttcttc	ttcttcttct	tgctgattgg	ttctataaga	aatctagtat	tttctttgaa	780
acagagtttt	cccgtggttt	tcgaacttgg	agaaagattg	ttaagcttct	gtatattctg	840
cccaaatttg	tcgggcccac	ggcgaaaaac	gttgcgattt	tcggcttatt	gttttctctt	900
cttgtgttgg	ttccttctca	gatcttcgct	gatcgaatct	gcactggaat	aacatcttca	960
aactcacctc	atgtggtcaa	aacagccact	caaggggagg	tcaatgtgac	tggtgtgata	1020
ccactaacia	caacaccaac	aaaatcttat	tttgcaaatac	tcaaaggaac	aaggaccaga	1080
gggaaactat	gccagactg	tctcaactgc	acagatctgg	atgtggcttt	gggcagacca	1140
atgtgtgtgg	ggaccacacc	ttcggcgaag	gcttcaatac	tccacgaagt	caaacctgtt	1200
acatccgggt	gctttcctat	aatgcacgac	agaacaaaaa	tcaggcaact	acccaatctt	1260
ctcagaggat	atgaaaatat	caggctatca	acccaaaacg	tcatcgatgc	ggaaaaggca	1320
ccaggaggac	cctacagact	tggaaacctca	ggatcttgcc	ctaacgctac	cagtaagagc	1380
ggatttttcg	caacaatggc	ttgggctgtc	ccaaaggaca	acaacaaaaa	tgcaacgaac	1440
ccactaacag	tagaagtacc	atacatttgt	acagaagggg	aagaccaaata	cactgttttg	1500
gggttccatt	cagataacia	aacccaaatg	aagaacctct	atggagactc	aaatcctcaa	1560
aagttcacct	catctgctaa	tggagtaacc	acacactatg	tttctcagat	tggcagcttc	1620
ccagatcaaa	cagaagacgg	aggactacca	caaagcggca	ggattgttgt	tgattacatg	1680
atgcaaaaac	ctgggaaaac	aggaacaatt	gtctacaaaa	gaggtgtttt	gttgacctcaa	1740
aaggtgtggt	gcgcgagtgg	caggagcaaa	gtaataaaaag	ggtccttgcc	tttaattggt	1800
gaagcagatt	gccttcatga	aaaatacggg	ggattaaaca	aaagcaagcc	ttactacaca	1860
ggagaacatg	caaaagccat	aggaaattgc	ccaatatggg	tgaaaacacc	tttgaagctc	1920
gccaatggaa	ccaaatatag	acctcctgca	aaactattaa	aggaaagggg	tttcttcgga	1980
gctattgctg	gtttcctaga	aggaggatgg	gaaggaatga	ttgcaggctg	gcacggatac	2040
acatctcacg	gagcacatgg	agtggcagtg	gcggcggacc	ttaagagtac	gcaagaagct	2100
ataaacaaga	taacaaaaaa	tctcaattct	ttgagtgagc	tagaagtaaa	gaatcttcaa	2160

agactaagtg gtgccatgga tgaactccac aacgaaatac tcgagctgga tgagaaagtg 2220
 gatgatctca gagctgacac tataagctcg caaatagaac ttgcagtctt gctttccaac 2280
 gaaggaataa taaacagtga agatgagcat ctattggcac ttgagagaaa actaaagaaa 2340
 atgctggggtc cctctgctgt agagatagga aatggatgct tcgaaaccaa acacaagtgc 2400
 aaccagacct gcttagacag gatagctgct ggcaccttta atgcaggaga attttctctc 2460
 cccacttttg attcactgaa cattactgct gcatctttta atgatgatgg attggataac 2520
 taccaaatac tgtcaattta ttcaacagtg gcgagttccc tagcactggc aatcatgatg 2580
 gctgggtctat ctttatggat gtgctccaat ggatcggttac aatgcagaat ttgcatttaa 2640
 aggcctatct tctttagttt gaatttactg ttattcgggtg tgcatttcta tgtttggtga 2700
 gcgggttttct gtgctcagag tgtgtttatt ttatgtaatt taatttcttt gtgagctcct 2760
 gtttagcagg tcgtcccttc agcaaggaca caaaaagatt ttaattttat taaaaaaaaa 2820
 aaaaaaaaaa accggaatt cgatatcaag cttatcgacc tgcagatcgt tcaaacattt 2880
 ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcatgatt atcatataat 2940
 ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg ttatttatga 3000
 gatgggtttt tatgattaga gtcccgaat tatacattta atacgcgata gaaaacaaaa 3060
 tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta ctagattcta 3120
 gagtctcaag cttcggcgcg cc 3142

<210> 76

<211> 1272

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự mã hóa Msj1 tổng hợp

<400> 76

atgtttgggc gcggaccaac aaggaagagt gataacacca aatattacga tattcttgggt 60
 gtttcaaaaa gtgctagtga agatgaaatc aagaaagcct atagaaaggc agcgatgaag 120
 aaccatccag ataaggggtg ggatcctgag aagttcaagg agttgggcca agcatatgaa 180
 gtgttgagcg atcctgaaaa gaaagaactg tatgatcaat atggtgaaga tgcccttaaa 240
 gaaggaatgg ggggaggcgc aggaagctca tttcataatc cgtttgatat tttcgaatca 300
 ttttttgggtg caggcttttg tgggtggtgg cttcacgcg caagaagaca gaagcaagga 360
 gaagatgtgg tgcattctat aaaggtttcc ttggaggatg tgtataacgg cactacaaag 420
 aagctatcac tttctaggaa tgcactgtgc tcaaatgta aagggaagg ttcaaaaagt 480
 ggaactgctg gaaggtgttt tggatgccag ggcacaggta tgaagattac cagaaggcaa 540

attggactgg gcatgattca acaaatgcaa cacgtctgtc ctgactgcaa aggaacaggc 600
 gaggtcatta gtgagagaga tagatgccct caatgcaagg gaaacaagat tactcaagaa 660
 aagaaggtgc tggaggtgca tgtggaaaag gggatgcagc agggtcacaa gattgtattc 720
 gaaggacaag ctgatgaagc tcctgataca atcacaggag acatagtttt tgtcttgcaa 780
 gtaaagggac atccgaagtt tcggagggag cgtgatgacc tccacattga acacaatttg 840
 agcttaactg aggctctctg tggcttccag tttaatgtca cacatcttga tgggaaggcaa 900
 ctattgggtca aatcgaaccc cggcgaagtc atcaagccag gtcaacataa agctataaat 960
 gatgagggaa tgccacaaca tggtagggccg ttcattgaagg gacgcctata catcaagttt 1020
 agtggttgatt tcccggattc gggttttctt tccccaaagg aaagcctgga attagaaaag 1080
 atattacctc aaaagacaag caagaacttg tccccaaagg aggtagatga ttgtgaggag 1140
 accaccctgc atgatgtcaa tattgcagag gagatgagtc gaaagaagca acaataccgt 1200
 gaggcataatg atgacgatga tgatgaagat gatgagcact cgcagcctcg ggtgcaatgc 1260
 gctcaacagt ag 1272

<210> 77
 <211> 4402
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Sản phẩm thiết kế số R850 tổng hợp, từ HindIII

<400> 77
 aagcttgcat gctgacaggt cgactctaga ggatccccgg gctgggtctgt acattcatct 60
 tgccgccttt gcattcactt ggccacaaag agtagagaga aggaagagaa gagcccagac 120
 ttcaagaagc gaccttgcaa gtgcactcga gggtcagaaa ctgtatatca tatctatgtg 180
 agagaaaggg gaacatttga gatggagtc atttacttga ggtatactta ttattttgat 240
 caataaattt gtatacttct tatttagatc aataaatttg tcattaagct ataatacaaa 300
 ataaattacg atcaaatatg caaatgttag ccagtacttg tgtaaaactt gatggcatct 360
 cttggtttct ttggcaatca catgcctaag aaataaatag tatcatatga ttgtgtttgg 420
 tcagacttca gagtcagatg actctgtttg gataaacagc ttaattaagc gcttatagaa 480
 tatcatatga ttgtgtttgg tcagacttca gagcatctct tggtttctct ggcaatcata 540
 tgcctaagaa ataaatagta tcatatgatt gtgtttgggtc agacttcaga gtcagatgac 600
 cctgtttggg taaacagctt aattaagtgc ttatagaata agcgcttata atataagtgc 660
 ttttgtacag ttatttctat gaaagtagaa gaaatagtca tattgtttta atataagcta 720
 tcctggagag cttgtggaaa taaccagaaa agaacttatg gacacgtcat gagctgttta 780

cataagatct	ccctaacagt	ctcaaaagt	gtg	tttatgccag	tagataaatt	caaataagtc	840	
aatctaaaca	gaccctaaat	ccattatggt	acctatcatt	ttagcttatt	ccatctttat		900	
taagaatgtc	atgagataac	ataatgataa	cacattat	ttt	tgacacaaat	gggcagatct	960	
agcaatttaa	ctctggagtc	cttcaagact	gctgttctta	cgaagttcac	gtccctgaat		1020	
catgttcctg	tatggaagcc	tgaaagacct	caaattctaa	aaggtggcga	taaattgaag		1080	
gtttacaaaa	tataccctgc	gggcttgaca	cagaggcaag	ctctttatac	cttcagttc		1140	
aacggggatg	ttgatttcag	aagtcacttg	gagagcaatc	cttgtgccaa	gtttgaagta		1200	
atTTTTgtgt	agcatatgtt	gagctaccta	caatttacat	gatcacctag	cattagctct		1260	
ttcacttaac	tgagagaatg	aagttttagg	aatgagtatg	accatggagt	cggcatggct		1320	
ttgtaatgcc	taccctactt	tgGCCaactc	atcggggatt	tacattcaga	aaatatacat		1380	
gacttcaacc	atacttaa	ac	ccctttttgt	aagataactg	aatgttcata	tttaatgttg	1440	
ggttgtagtg	tttttacttg	attatatcca	gacagttaca	agttggacaa	caagattgtg		1500	
ggtctgtact	gttatttatt	tatttttttt	ttagcagaaa	caccttatct	tttgtttcgt		1560	
ttgaatgtag	aatgaaaata	aaagaaagaa	aatataacat	catcgccgc	gcttgtctaa		1620	
tttcgggcag	ttaggatcct	ctccgggtcac	cggaaagttt	cagtagaaga	aacaaaacac		1680	
cgtgactaaa	atgatactat	tattttat	ttt	atttgtgttt	tcttttttct	accggaactt	1740	
tttagaacgg	atcccaactc	gttccggggc	cgctacaact	gaaacaaaag	aagatat	ttt	1800	
ctctctcttc	agaaatgtaa	gttttccttt	acagataccc	attcaccatt	tgattcagat		1860	
gtggtgacta	gagataaagc	atactaattt	gactcttgga	aaccataaa	gtttatgtta		1920	
tccgtgttct	ggaccaatcc	acttgggggc	ataacctgtg	tctatgtgtg	gtttggtttc		1980	
cattctgatt	tatgcggcga	cttgtaattt	aaaatctagg	aggggcagac	attgaacaat		2040	
cccaatat	ttt	taataactta	tgcaagattt	tttttattaa	tgagatgatg	tgtttgtagc	2100	
tgagattgag	tcatacat	ttt	cactaagaaa	tggttccaag	taccaaacta	tcatgaccca	2160	
gttgcaaaca	tgacgttcgg	gagtggtcac	tttgatagtt	caatttcac	ttggcttctt		2220	
attcctttta	taattcta	att	tcttcttg	tg	ttaa	catgtattat	tttcttttaa	2280
aatttacatg	tcatttat	ttt	tgctcacta	actcaatt	ttt	gcatataaca	atgataagtg	2340
atattttgac	tcacaaaatt	tacatcaa	att	tcgacatcg	tttattatgt	tcattggatg		2400
attaacaaat	ataacaaact	ttgcaactaa	ttaaccacca	actgaatata	attaactata			2460
actgtgaaag	tagttaactc	at	tttttata	ttcatagatc	aaataagaga	aataacggta		2520
tattaatccc	tccaaaaaaa	aaaaacggta	tattttactaa	aaaatctaag	ccacgtagga			2580
ggataacagg	atccccgtag	gaggataaca	tccaatccaa	ccaatcacia	caatcctgat			2640

gagataaccc	actttaagcc	cacgcatctg	tggcacatct	acattatcta	aatcacacat	2700
tcttccacac	atctgagcca	cacaaaaacc	aatccacatc	tttatcacc	attctataaa	2760
aatcacact	ttgtgagtct	acactttgat	tcccttcaaa	cacatacaaa	gagaagagac	2820
taattaatta	attaatcatc	ttgagagaaa	atgtttgggc	gcggaccaac	aaggaagagt	2880
gataacacca	aatattacga	tattcttggg	gtttcaaaaa	gtgctagtga	agatgaaatc	2940
aagaaagcct	atagaaaggc	agcgatgaag	aaccatccag	ataaggggtg	ggatcctgag	3000
aagttcaagg	agttgggcca	agcatatgaa	gtgttgagcg	atcctgaaaa	gaaagaactg	3060
tatgatcaat	atgggtgaaga	tgcccttaaa	gaaggaatgg	ggggaggcgc	aggaagctca	3120
tttcataatc	cgtttgatat	tttcgaatca	ttttttgggtg	caggctttgg	tggtgggtgg	3180
ccttcacgcg	caagaagaca	gaagcaagga	gaagatgtgg	tgcattctat	aaaggtttcc	3240
ttggaggatg	tgtataacgg	cactacaaa	aagctatcac	tttctaggaa	tgcactgtgc	3300
tcaaaatgta	aagggaagg	ttcaaaaagt	ggaactgctg	gaaggtgttt	tggatgccag	3360
ggcacaggta	tgaagattac	cagaaggcaa	attggactgg	gcatgattca	acaaatgcaa	3420
cacgtctgtc	ctgactgcaa	aggaacaggc	gaggtcatta	gtgagagaga	tagatgccct	3480
caatgcaagg	gaaacaagat	tactcaagaa	aagaaggtgc	tggaggtgca	tgtggaaaag	3540
gggatgcagc	agggtcacia	gattgtattc	gaaggacaag	ctgatgaagc	tcctgatata	3600
atcacaggag	acatagtttt	tgtcttgcaa	gtaaaggggac	atccgaagtt	tcggagggag	3660
cgtgatgacc	tccacattga	acacaatttg	agcttaactg	aggctctctg	tggtctccag	3720
tttaatgtca	cacatcttga	tggaaggcaa	ctattgggtca	aatcgaaccc	cggcgaagtc	3780
atcaagccag	gtcaacataa	agctataaat	gatgagggaa	tgccacaaca	tggtaggccg	3840
ttcatgaagg	gacgcctata	catcaagttt	agtgttgatt	tcccggattc	gggttttctt	3900
tccccaagcc	aaagcctgga	attagaaaag	atattacctc	aaaagacaag	caagaacttg	3960
tcccaaaagg	aggtagatga	ttgtgaggag	accaccctgc	atgatgtcaa	tattgcagag	4020
gagatgagtc	gaaagaagca	acaataccgt	gaggcatatg	atgacgatga	tgatgaagat	4080
gatgagcact	cgcagcctcg	ggtgcaatgc	gctcaacagt	aggagctcag	ctcgaatttc	4140
cccgatcggt	caaacatttg	gcaataaagt	ttcttaagat	tgaatcctgt	tgccggtctt	4200
gcgatgatta	tcatataatt	tctgttgaat	tacgttaagc	atgtaataat	taacatgtaa	4260
tgcatgacgt	tatttatgag	atgggttttt	atgattagag	tcccgcaatt	atacatttaa	4320
tacgcgatag	aaaacaaaat	atagcgcgca	aactaggata	aattatcgcg	cgcggtgtca	4380
tctatgttac	tagatcgaat	tc				4402

<210> 78
 <211> 5086
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Sản phẩm thiết kế số R860 tổng hợp, từ HindIII đến EcoRI

<400> 78
 aagcttgcac gcctgcaggt cgactctaga ggatccccgg gctgggtctgt acattcatct 60
 tgccgccttt gcattcactt ggccacaaag agtagagaga aggaagagaa gagcccagac 120
 ttcaagaagc gaccttgcaa gtgcactcga gggtcagaaa ctgtatatca tatctatgtg 180
 agagaaaggg gaacatttga gatggagtcc atttacttga ggtataactta ttattttgat 240
 caataaattt gtatacttct tatttagatc aataaatttg tcattaagct ataatccaaa 300
 ataaattacg atcaaatatg caaatgttag ccagtacttg tgttaaactt gatggcatct 360
 cttggtttct ttggcaatca catgcctaag aaataaatag tatcatatga ttgtgtttgg 420
 tcagacttca gagtcagatg actctgtttg gataaacagc ttaattaagc gcttatagaa 480
 tatcatatga ttgtgtttgg tcagacttca gagcatctct tggtttctct ggcaatcata 540
 tgcctaagaa ataaatagta tcatatgatt gtgtttgggtc agacttcaga gtcagatgac 600
 cctgtttggg taaacagctt aattaagtgc ttatagaata agcgcttctc atataagtgc 660
 ttttgtacag ttatttctat gaaagtagaa gaaatagtca tattgtttta atataagcta 720
 tcctggagag cttgtggaaa taaccagaaa agaacttatg gacacgtcat gagctgttta 780
 cataagatct ccctaacagt ctcaaaagtg tttatgccag tagataaatt caaataagtc 840
 aatctaaaca gaccctaaat ccattatggg acctatcatt ttagcttatt ccatctttat 900
 taagaatgtc atgagataac ataatagata cacattatct tgacacaaat gggcagatct 960
 agcaatttaa ctctggagtc cttcaagact gctgtttctta cgaagtccac gtccctgaat 1020
 catgttcctg tatggaagcc tgaaagacct caaattctaa aaggtggcga taaattgaag 1080
 gttttacaaaa tataccctgc gggcttgaca cagaggcaag ctctttatac cttccagttc 1140
 aacgggggatg ttgatttcag aagtcacttg gagagcaatc cttgtgcaa gtttgaagta 1200
 atttttgtgt agcatatgtt gagctaccta caatttacat gatcacctag cattagctct 1260
 ttacttaac tgagagaatg aagttttagg aatgagtatg accatggagt cggcatggct 1320
 ttgtaatgcc taccctactt tggccaactc atcgggggatt tacattcaga aaatatacat 1380
 gacttcaacc atacttaaac ccctttttgt aagataactg aatgttcata tttaatgttg 1440
 gggtgtagtg tttttacttg attatatcca gacagttaca agttggacaa caagattgtg 1500
 ggtctgtact gttattttatt tttttttttt ttagcagaaa caccttatct tttgtttcgt 1560

ttgaatgtag	aatgaaaata	aaagaaagaa	aatataacat	catcggccgc	gcttgtctaa	1620
tttcgggcag	ttaggatcct	ctccggtcac	cggaagttt	cagtagaaga	aacaaaacac	1680
cgtgactaaa	atgatactat	tattttat	attgtgtttt	tcttttttct	accggaactt	1740
tttagaacgg	atcccaactc	gttccggggc	cgctacaact	gaaacaaaag	aagatat	1800
ctctctcttc	agaaatgtaa	gttttccttt	acagataccc	attcaccatt	tgattcagat	1860
gtggtgacta	gagataaagc	atactaattt	gactccttga	aaccataaa	gtttatgtta	1920
tccgtgttct	ggaccaatcc	acttgggggc	ataacctgtg	tctatgtgtg	gtttggtttc	1980
cattctgatt	tatgcggcga	cttghtaattt	aaaatctagg	aggggcagac	attgaacaat	2040
cccaatat	taataactta	tgcaagattt	tttttattaa	tgagatgatg	tgtttgtgac	2100
tgagattgag	tcatacat	cactaagaaa	tggttccaag	taccaaacta	tcatgaccca	2160
gttgcaaaca	tgacgttcgg	gagtggtcac	ttgatagtt	caatttcac	ttggcttctt	2220
attcctttta	taattcta	tcttctgtg	taaactat	catgtattat	tttcttttaa	2280
aatttacatg	tcatttat	tgctcacta	actcaatt	gcatataaca	atgataagtg	2340
atattttgac	tcacaaaatt	tacatcaa	ttcgacatcg	tttattatgt	tcattggatg	2400
attaacaaat	ataacaaact	ttgcaactaa	ttaaccacca	actgaatata	attaactata	2460
actgtgaaag	tagttaactc	atttttatat	ttcatagatc	aaataagaga	aataacggta	2520
tattaatccc	tcaaaaaaaa	aaaaacggta	tatttactaa	aaaatctaag	ccacgtagga	2580
ggataacagg	atccccgtag	gaggataaca	tccaatccaa	ccaatcacia	caatcctgat	2640
gagataaccc	actttaagcc	cacgcatctg	tggcacatct	acattatcta	aatcacacat	2700
tcttccacac	atctgagcca	cacaaaaacc	aatccacatc	tttatcacc	attctataaa	2760
aatcacact	ttgtgagtct	acactttgat	tcccttcaaa	cacatacaaa	gagaagagac	2820
taattaatta	attaatcatc	ttgagagaaa	atgtcgggta	aaggagaagg	accagctatc	2880
ggtatcgatc	ttggtaccac	ttactcttgc	gtcggagtat	ggcaacacga	ccgtgttgag	2940
atcattgcta	atgatcaagg	aaacagaacc	acgccatctt	acgttgcttt	caccgactcc	3000
gagaggttga	tcggtgacgc	agctaagaat	caggtcgcca	tgaaccccg	taacaccgtt	3060
ttcgacgcta	agaggttgat	cggtcgtcgt	ttctctgaca	gctctgttca	gagtgacatg	3120
aaattgtggc	cattcaagat	tcaagccgga	cctgccgata	agccaatgat	ctacgtcgaa	3180
tacaaggggtg	aagagaaaga	gttcgcagct	gaggagattt	cttccatgg	tcttattaag	3240
atgcgtgaga	ttgctgaggc	ttaccttgg	gtcacaaatca	agaacgccgt	tgttaccgtt	3300
ccagcttact	tcaacgactc	tcagcgtcag	gctacaaagg	atgctgggtg	catcgctgg	3360
ttgaacgtta	tgcgaaatcat	caacgagcct	acagccgccg	ctattgccta	cggtcttgac	3420

```

aaaaaggcta ccagcgttgg agagaagaat gttcttatct tcgatcttgg tgggtggcact 3480
tttgatgtct ctcttcttac cattgaagag ggtatctttg aggtgaaggc aactgctggt 3540
gacacccatc ttggtgggga agattttgac aacagaatgg ttaaccactt tgtccaagag 3600
ttcaagagga agagtaagaa ggatatcacc ggtaacccaa gagctcttag gaggttgaga 3660
acttcctgtg agagagcgaa gaggactctt tcttccactg ctcagaccac catcgagatt 3720
gactctctat acgaggggtat cgacttctac tccaccatca cccgtgctag atttgaggag 3780
ctcaacatgg atctcttcag gaagtgtatg gagccagttg agaagtgtct tcgtgatgct 3840
aagatggaca agagcactgt tcatgatgtt gtccttgttg gtggttctac ccgatatccct 3900
aaggttcagc aattgctcca ggacttcttc aacggcaaag agctttgcaa gtctattaac 3960
cctgatgagg ctggtgccta cggtgctgct gtccaggagg ctattctcag cgggtgaagga 4020
aacgagaagg ttcaagatct tctattgctc gatgtcactc ctctctccct tggtttgaa 4080
actgccggtg gtgtcatgac cactttgatc ccaaggaaca caaccatccc aaccaagaag 4140
gaacaagtct tctccaccta ctcagacaac caaccgggtg tgttgatcca ggtgtacgaa 4200
ggagagagag ccagaaccaa ggacaacaac cttcttggtg aatttgagct ctccggaatt 4260
cctccagctc ctcgtggtgt ccccagatc acagtctgct ttgacattga tgccaatggt 4320
atcctcaatg tctctgctga ggacaagacc accggacaga agaacaagat caccatcacc 4380
aatgacaagg gtcgtctctc caaggatgag attgagaaga tggttcaaga ggctgagaag 4440
tacaagtccg aagacgagga gcacaagaag aaggttgaag ccaagaacgc tctcgagaac 4500
tacgcttaca acatgaggaa caccatccaa gacgagaaga ttggtgagaa gctcccggct 4560
gcagacaaga agaagatcga ggattctatt gagcaggcga ttcaatggct cgagggtaac 4620
cagttggctg aggctgatga gttcgaagac aagatgaagg aattggagag catctgcaac 4680
ccaatcattg ccaagatgta ccaaggagct ggtggtgaag ccggtggtcc aggtgcctct 4740
ggtatggacg atgatgctcc ccctgcttca ggcggtgctg gacctaagat cgaggaggtc 4800
gactaagagc tcagctcgaa tttccccgat cgttcaaaca ttgggcaata aagtttctta 4860
agattgaatc ctggtgccgg tcttgcgatg attatcatat aatttctggt gaattacgtt 4920
aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt 4980
agagtcccg c aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag 5040
gataaattat cgcgcgcggt gtcattctatg ttactagatc gaattc 5086

```

<210> 79

<211> 9493

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế số R870 tổng hợp, từ HindIII đến EcoRI

<400> 79

```

aagcttgcat gcctgcaggt cgactctaga ggatccccgg gctgggtctgt acattcatct      60
tgccgccttt gcattcactt ggccacaaag agtagagaga aggaagagaa gagcccagac      120
ttcaagaagc gaccttgcaa gtgcactcga gggtcagaaa ctgtatatca tatctatgtg      180
agagaaaggg gaacatttga gatggagtcc atttacttga ggtatactta ttattttgat      240
caataaattt gtatacttct tatttagatc aataaatttg tcattaagct ataatccaaa      300
ataaattacg atcaaatatg caaatgttag ccagtacttg tgttaaactt gatggcatct      360
cttggtttct ttggcaatca catgcctaag aaataaatag tatcatatga ttgtgtttgg      420
tcagacttca gagtcagatg actctgtttg gataaacagc ttaattaagc gcttatagaa      480
tatcatatga ttgtgtttgg tcagacttca gagcatctct tggtttctct ggcaatcata      540
tgcctaagaa ataaatagta tcatatgatt gtgtttgggtc agacttcaga gtcagatgac      600
cctgtttggg taaacagctt aattaagtgc ttatagaata agcgcttatt atataagtgc      660
ttttgtacag ttatctctat gaaagtagaa gaaatagtca tattgtttta atataagcta      720
tcctggagag cttgtggaaa taaccagaaa agaacttatg gacacgtcat gagctgttta      780
cataagatct ccctaacagt ctcaaaagtg tttatgccag tagataaatt caaataagtc      840
aatctaaaca gaccctaaat ccattatggg acctatcatt ttagcttatt ccatctttat      900
taagaatgtc atgagataac ataatagata cacattatct tgacacaaat gggcagatct      960
agcaatttaa ctctggagtc cttcaagact gctgttctta cgaagttcac gtccctgaat     1020
catgttcctg tatggaagcc tgaaagacct caaattctaa aagggtggcg taaattgaag     1080
gtttacaaaa tataccctgc gggcttgaca cagaggcaag ctctttatac cttccagttc     1140
aacgggggatg ttgatttcag aagtcacttg gagagcaatc cttgtgcca gtttgaagta     1200
atTTTTgtgt agcatatggt gagctaccta caatttacat gatcacctag cattagctct     1260
ttcacttaac tgagagaatg aagttttagg aatgagtatg accatggagt cggcatggct     1320
ttgtaatgcc taccctactt tggccaactc atcgggggatt tacattcaga aaatatacat     1380
gacttcaacc atacttaaac ccctttttgt aagataactg aatgttcata tttaatgttg     1440
ggttgtagtg tttttacttg attatatcca gacagttaca agttggacaa caagattgtg     1500
ggctgtgact gttatttatt tttttttttt ttagcagaaa caccttatct tttgtttcgt     1560
ttgaatgtag aatgaaaata aaagaaagaa aatataacat catcggccgc gcttgtctaa     1620
tttcggggcag ttaggatcct ctccgggtcac cggaaagttt cagtagaaga aacaaaacac     1680

```

cgtgactaaa atgatactat tatttttattt attgtgtttt tcttttttct accggaactt 1740
 tttagaacgg atcccaactc gttccggggc cgctacaact gaaacaaaag aagatatattt 1800
 ctctctcttc agaaatgtaa gttttccttt acagataccc attcaccatt tgattcagat 1860
 gtggtgacta gagataaagc atactaattt gactccttga aacccataaa gtttatgtta 1920
 tccgtgttct ggaccaatcc acttgggggc ataacctgtg tctatgtgtg gtttggtttc 1980
 cattctgatt tatgcggcga cttgtaattt aaaatctagg aggggcagac attgaacaat 2040
 cccaatatatt taataactta tgcaagattt tttttattaa tgagatgatg tgtttgtgac 2100
 tgagattgag tcatacattt cactaagaaa tggttccaag taccaaacta tcatgaccca 2160
 gttgcaaaca tgacgttcgg gagtggtcac tttgatagtt caatttcac tggccttctt 2220
 attcctttta taattctaatt tcttcttgtg taaactattt catgtattat ttttctttta 2280
 aatttacatg tcatttatatt tgcctcacta actcaatttt gcatataaca atgataagtg 2340
 atattttgac tcacaaaatt tacatcaaatt ttcgacatcg tttattatgt tcattggatg 2400
 attaacaaat ataacaaact ttgcaactaa ttaaccacca actgaatata attaactata 2460
 actgtgaaag tagttaactc atttttatat ttcatagatc aaataagaga aataacggta 2520
 tattaatccc tccaaaaaaa aaaaacggta tatttactaa aaaatctaag ccacgtagga 2580
 ggataacagg atccccgtag gaggataaca tccaatccaa ccaatcaca caatcctgat 2640
 gagataaccc actttaagcc cacgcatctg tggcacatct acattatcta aatcacacat 2700
 tcttccacac atctgagcca cacaaaaacc aatccacatc tttatcacc attctataaa 2760
 aatcacact ttgtgagtct acactttgat tcccttcaaa cacatacaaa gagaagagac 2820
 taattaatta attaatcatc ttgagagaaa atgtcgggta aaggagaagg accagctatc 2880
 ggtatcgatc ttggtaccac ttactcttgc gtcggagtat ggcaacacga ccgtgttgag 2940
 atcattgcta atgatcaagg aaacagaacc acgccatctt acgttgcttt caccgactcc 3000
 gagaggttga tcggtgacgc agctaagaat caggtcgcca tgaacccgt taacaccgtt 3060
 ttcgacgcta agaggttgat cggtcgtcgt ttctctgaca gctctgttca gagtgacatg 3120
 aaattgtggc cattcaagat tcaagccgga cctgccgata agccaatgat ctacgtcgaa 3180
 tacaaggggtg aagagaaaga gttcgcagct gaggagattt cttccatggt tcttattaag 3240
 atgcgtgaga ttgctgaggc ttaccttggg gtcacaatca agaacgccgt tgttaccgtt 3300
 ccagcttact tcaacgactc tcagcgtcag gctacaaagg atgctgggtg catcgctggg 3360
 ttgaacgtta tgcgaatcat caacgagcct acagccgccg ctattgccta cggctcttgac 3420
 aaaaaggcta ccagcgttgg agagaagaat gttcttatct tcgatcttgg tgggtggcact 3480
 tttgatgtct ctcttcttac cattgaagag ggtatctttg aggtgaaggc aactgctggt 3540

gacacccatc ttggtgggga agattttgac aacagaatgg ttaaccactt tgtccaagag 3600
ttcaagagga agagtaagaa ggatatcacc ggtaacccaa gagctcttag gaggttgaga 3660
acttctgtg agagagcgaa gaggactctt tcttccactg ctcagaccac catcgagatt 3720
gactctctat acgaggggtat cgacttctac tccaccatca cccgtgctag atttgaggag 3780
ctcaacatgg atctcttcag gaagtgtatg gagccagttg agaagtgtct tcgtgatgct 3840
aagatggaca agagcactgt tcatgatgtt gtccttggtg gtggttctac ccgtatccct 3900
aaggttcagc aattgctcca ggacttcttc aacggcaaag agctttgcaa gtctattaac 3960
cctgatgagg ctggtgccta cggtgctgct gtccaggag ctattctcag cgggtgaagga 4020
aacgagaagg ttcaagatct tctattgctc gatgtcactc ctctctccct tggtttgaa 4080
actgccgtg gtgtcatgac cactttgatc ccaaggaaca caaccatccc aaccaagaag 4140
gaacaagtct tctccaccta ctcagacaac caacccggtg tgttgatcca ggtgtacgaa 4200
ggagagagag ccagaaccaa ggacaacaac cttcttggtg aatttgagct ctccggaatt 4260
cctccagctc ctctggtgt ccccagatc acagtctgct ttgacattga tgccaatggt 4320
atcctcaatg tctctgctga ggacaagacc accggacaga agaacaagat caccatcacc 4380
aatgacaagg gtcgtctctc caaggatgag attgagaaga tggttcaaga ggctgagaag 4440
tacaagtccg aagacgagga gcacaagaag aaggttgaag ccaagaacgc tctcgagaac 4500
tacgcttaca acatgaggaa caccatcaa gacgagaaga ttggtgagaa gctcccggct 4560
gcagacaaga agaagatcga ggattctatt gagcaggcga ttcaatggct cgagggtaac 4620
cagttggctg aggctgatga gttcgaagac aagatgaagg aattggagag catctgcaac 4680
ccaatcattg ccaagatgta ccaaggagct ggtggtgaag ccggtggtcc aggtgcctct 4740
ggtatggacg atgatgctcc ccctgcttca ggcggtgctg gacctaagat cgaggaggtc 4800
gactaagagc tcagctcgaa tttccccgat cgttcaaaca tttggcaata aagtttctta 4860
agattgaatc ctggtgccgg tcttgcgatg attatcatat aatttctggt gaattacgtt 4920
aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt ttttatgatt 4980
agagtccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag 5040
gataaattat cgcgcgcggt gtcattctatg ttactagatc gaattcgtaa tcatggctcat 5100
agctgtttcc tgtgtgaaat tgttatccgg ggctgggtctg tacattcatc ttgccgcctt 5160
tgcattcact tggccacaaa gagtagagag aaggaagaga agagcccaga cttcaagaag 5220
cgaccttgca agtgcactcg agggtcagaa actgtatatc atatctatgt gagagaaagg 5280
ggaacatttg agatggagtc catttacttg aggtatactt attattttga tcaataaatt 5340

tgtatacttc	ttatttagat	caataaattt	gtcattaagc	tataatccaa	aataaattac	5400
gatcaaatat	gcaaatgtta	gccagtactt	gtgttaaact	tgatggcatc	tcttggtttc	5460
tttggaatc	acatgcctaa	gaaataaata	gtatcatatg	attgtgtttg	gtcagacttc	5520
agagtcagat	gactctgttt	ggataaacag	cttaattaag	cgcttataga	atatcatatg	5580
attgtgtttg	gtcagacttc	agagcatctc	ttggtttctc	tggcaatcat	atgcctaaga	5640
aataaatagt	atcatatgat	tgtgtttggg	cagacttcag	agtcagatga	ccctgtttgg	5700
gtaaacagct	taattaagtg	cttatagaat	aagcgcttat	catataagtg	cttttgtaca	5760
gttattttcta	tgaaagtaga	agaaatagtc	atattgtttt	aatataagct	atcctggaga	5820
gcttgtggaa	ataaccagaa	aagaacttat	ggacacgtca	tgagctgttt	acataagatc	5880
tccctaacag	tctcaaaagt	gtttatgcca	gtagataaat	tcaaataagt	caatctaaac	5940
agaccctaaa	tccattatgg	tacctatcat	tttagcttat	tccatcttta	ttaagaatgt	6000
catgagataa	cataatgata	acacattatt	ttgacacaaa	tgggcagatc	tagcaattta	6060
actctggagt	ccttcaagac	tgctgttctt	acgaagtcca	cgccctgaa	tcatgttcct	6120
gtatggaagc	ctgaaagacc	tcaaattcta	aaagggtggc	ataaattgaa	ggtttacaaa	6180
atataccctg	cgggcttgac	acagaggcaa	gctctttata	ccttcagtt	caacggggat	6240
gttgatttca	gaagtcactt	ggagagcaat	ccttgtgcca	agtttgaagt	aatttttgtg	6300
tagcatatgt	tgagctacct	acaatttaca	tgatcaccta	gcattagctc	tttcaactaa	6360
ctgagagaat	gaagttttag	gaatgagtat	gaccatggag	tcggcatggc	tttgtaatgc	6420
ctaccctact	ttggccaact	catcggggat	ttacattcag	aaaatataca	tgacttcaac	6480
cataacttaa	cccctttttg	taagataact	gaatgttcat	atttaatgtt	gggttgtagt	6540
gtttttactt	gattatatcc	agacagttac	aagttggaca	acaagattgt	gggtctgtac	6600
tgttatttat	ttattttttt	tttagcagaa	acaccttatc	ttttgtttcg	tttgaatgta	6660
gaatgaaaat	aaaagaaaga	aaatataaca	tcacgggccg	cgcttgtcta	atttcgggca	6720
gttaggatcc	tctccggtca	ccggaaagt	tcagtagaag	aaacaaaaca	ccgtgactaa	6780
aatgatacta	ttattttatt	tattgtgttt	ttcttttttc	taccggaact	ttttagaacg	6840
gatcccaact	cgttccgggg	ccgctacaac	tgaacaaaaa	gaagatattt	tctctctctt	6900
cagaaatgta	agttttcctt	tacagatacc	cattcaccat	ttgattcaga	tgtggtgact	6960
agagataaag	catactaatt	tgactcttgg	aaaccataa	agtttatgtt	atccgtgttc	7020
tggaccaatc	cacttggggg	cataacctgt	gtctatgtgt	ggtttggttt	ccattctgat	7080
ttatgcggcg	acttgtaatt	taaaatctag	gaggggcaga	cattgaacaa	tcccaatatt	7140
ttaataactt	atgcaagatt	ttttttatta	atgagatgat	gtgtttgtga	ctgagattga	7200

gtcatacatt	tcactaagaa	atggttccaa	gtaccaaact	atcatgaccc	agttgcaaac	7260
atgacgttcg	ggagtgggtca	ctttgatagt	tcaatttcat	cttggcttct	tattcctttt	7320
ataattctaa	ttcttcttgt	gtaaactatt	tcatgtatta	tttttcttta	aaattttacat	7380
gtcattttatt	ttgcctcact	aactcaattt	tgcataatac	aatgataagt	gatatttttga	7440
ctcacaaaaat	ttacatcaaa	tttcgacatc	gtttattatg	ttcattggat	gattaacaaa	7500
tataacaaaac	tttgcaacta	attaaccacc	aactgaatat	aattaactat	aactgtgaaa	7560
gtagttaact	cattttttata	tttcatagat	caaataagag	aaataacggt	atattaatcc	7620
ctccaaaaaa	aaaaaacggt	atatttacta	aaaaatctaa	gccacgtagg	aggataacag	7680
gatccccgta	ggaggataac	atccaatcca	accaatcaca	acaatcctga	tgagataacc	7740
cactttaagc	ccacgcatct	gtggcacatc	tacattatct	aatcacaca	ttcttccaca	7800
catctgagcc	acacaaaaac	caatccacat	ctttatcacc	cattctataa	aaaatcacac	7860
tttgtgagtc	tacactttga	ttcccttcaa	acacatacaa	agagaagaga	ctaattaatt	7920
aattaatcat	cttgagagaa	aatgtttggg	cgcggaacca	caaggaagag	tgataacacc	7980
aaatattacg	atattcttgg	tgtttcaaaa	agtgtctagt	aagatgaaat	caagaaagcc	8040
tatagaaagg	cagcgatgaa	gaaccatcca	gataaggggtg	gggatcctga	gaagttcaag	8100
gagttggggc	aagcatatga	agtgttgagc	gatcctgaaa	agaaagaact	gtatgatcaa	8160
tatgggtgaag	atgcccttaa	agaaggaatg	gggggaggcg	caggaagctc	atttcataat	8220
ccgtttgata	ttttcgaatc	attttttggg	gcaggctttg	gtgggtggtg	tccttcacgc	8280
gcaagaagac	agaagcaagg	agaagatgtg	gtgcattcta	taaaggtttc	cttgagggat	8340
gtgtataacg	gcactacaaa	gaagctatca	ctttctagga	atgcactgtg	ctcaaaatgt	8400
aaagggaaag	gttcaaaaag	tggaactgct	ggaagggtgt	ttggatgcca	gggcacaggt	8460
atgaagatta	ccagaaggca	aattggactg	ggcatgattc	aacaaatgca	acacgtctgt	8520
cctgactgca	aaggaacagg	cgaggtcatt	agtgaagag	atagatgcc	tcaatgcaag	8580
ggaacaaga	ttactcaaga	aaagaagggtg	ctggagggtg	atgtggaaaa	ggggatgcag	8640
cagggtcaca	agattgtatt	cgaaggacaa	gctgatgaag	ctcctgatac	aatcacagga	8700
gacatagttt	ttgtcttgca	agtaaaggga	catccgaagt	ttcggaggga	gcgtgatgac	8760
ctccacattg	aacacaattt	gagcttaact	gaggctctct	gtggcttcca	gtttaatgtc	8820
acacatcttg	atggaaggca	actattgggtc	aaatcgaacc	ccggcgaagt	catcaagcca	8880
ggtcaacata	aagctataaa	tgatgaggga	atgccacaac	atggtaggcc	gttcatgaag	8940
ggacgcctat	acatcaagtt	tagtgttgat	ttcccggatt	cgggttttct	ttccccaagc	9000

```

caaagcctgg aattagaaaa gatattacct caaaagacaa gcaagaactt gtcccaaaag 9060
gaggtagatg attgtgagga gaccaccctg catgatgtca atattgcaga ggagatgagt 9120
cgaaagaagc aacaataccg tgaggcatat gatgacgatg atgatgaaga tgatgagcac 9180
tcgcagcctc ggggtgcaatg cgctcaacag taggagctca gctcgaattt ccccgatcgt 9240
tcaaacattht ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt 9300
atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg 9360
ttattttatga gatgggtttt tatgattaga gtccccgaat tatacattta atacgcgata 9420
gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc atctatgtta 9480
ctagatcgaa ttc 9493

```

<210> 80
 <211> 568
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Sản phẩm dịch mã axit amin tổng hợp của trình tự mã hóa trong
 catxet biểu hiện sản phẩm thiết kế 690

<400> 80

```

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 His Phe Ala
                100                      105                      110

Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
        115                      120                      125

Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val
        130                      135                      140

Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr
145                      150                      155                      160

Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu
                165                      170                      175

```

Ser Lys Ser Tyr Ala Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp
 180 185 190
 Gly Val His His Pro Pro Asn Ile Gly Asp Gln Lys Ala Leu Tyr His
 195 200 205
 Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Lys
 210 215 220
 Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly
 225 230 235 240
 Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile
 245 250 255
 Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu
 260 265 270
 Ser Arg Gly Phe Gly Ser Gly 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> 81

<211> 568

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm dịch mã axit amin tổng hợp của trình tự mã hóa trong catxet biểu hiện sản phẩm thiết kế 691

<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 Leu Leu Lys Gly Ile Ala
50 55 60

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

Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile Val
85 90 95

Glu Lys Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala
100 105 110

Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
115 120 125

Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val
130 135 140

Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr
145 150 155 160

Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu
165 170 175

Ser Lys Ser Tyr Ala Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp
180 185 190

Gly 195	Val	His	His	Pro	Pro	Asn	Ile	Gly	Asp	Gln	Lys	Ala	Leu	Tyr	His
Thr 210	Glu	Asn	Ala	Tyr	Val	Ser	Val	Val	Ser	Ser	His	Tyr	Ser	Arg	Lys
Phe 225	Thr	Pro	Glu	Ile	Ala	Lys	Arg	Pro	Lys	Val	Arg	Asp	Gln	Glu	Gly
Arg	Ile	Asn	Tyr	Tyr	Trp	Thr	Leu	Leu	Glu	Pro	Gly	Asp	Thr	Ile	Ile
Phe	Glu	Ala	Asn	Gly	Asn	Leu	Ile	Ala	Pro	Arg	Tyr	Ala	Phe	Ala	Leu
Ser	Arg	Gly	Phe	Gly	Ser	Gly	Ile	Ile	Asn	Ser	Asn	Ala	Pro	Met	Asp
Lys	Cys	Asn	Thr	Lys	Cys	Gln	Thr	Pro	Met	Gly	Ala	Ile	Asn	Ser	Ser
Met 305	Pro	Phe	His	Asn	Ile	His	Pro	Leu	Thr	Ile	Gly	Glu	Cys	Pro	Lys
Tyr	Val	Lys	Ser	Asn	Arg	Leu	Val	Leu	Ala	Thr	Gly	Leu	Arg	Asn	Ser
Pro	Gln	Arg	Glu	Ser	Arg	Arg	Lys	Lys	Arg	Gly	Leu	Phe	Gly	Ala	Ile
Ala	Gly	Phe	Ile	Glu	Gly	Gly	Trp	Gln	Gly	Met	Val	Asp	Gly	Trp	Tyr
Gly	Tyr	His	His	Ser	Asn	Glu	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Lys
Glu 385	Ser	Thr	Gln	Lys	Ala	Ile	Asp	Gly	Val	Thr	Asn	Lys	Val	Asn	Ser
Ile	Ile	Asp	Lys	Met	Asn	Thr	Gln	Phe	Glu	Ala	Val	Gly	Arg	Glu	Phe
Asn	Asn	Leu	Glu	Arg	Arg	Ile	Glu	Asn	Leu	Asn	Lys	Lys	Met	Glu	Asp
Gly	Phe	Leu	Asp	Val	Trp	Thr	Tyr	Asn	Ala	Glu	Leu	Leu	Val	Leu	Met
Glu	Asn	Glu	Arg	Thr	Leu	Asp	Phe	His	Asp	Ser	Asn	Val	Lys	Asn	Leu
Tyr 465	Asp	Lys	Val	Arg	Leu	Gln	Leu	Arg	Asp	Asn	Ala	Lys	Glu	Leu	Gly
Asn	Gly	Cys	Phe	Glu	Phe	Tyr	His	Lys	Cys	Asp	Asn	Glu	Cys	Met	Glu
Ser	Ile	Arg	Asn	Gly	Thr	Tyr	Asn	Tyr	Pro	Gln	Tyr	Ser	Glu	Glu	Ala

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> 82

<211> 572

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm dịch mã axit amin tổng hợp của trình tự mã hóa trong catxet biểu hiện sản phẩm thiết kế 696

<400> 82

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 Thr Ile Cys Ile Gly Tyr His
20 25 30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
35 40 45

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

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65 70 75 80

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
85 90 95

Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
100 105 110

Thr Cys Tyr Pro Gly Ser Phe Asn Asp Tyr Glu Glu Leu Lys His Leu
115 120 125

Leu Ser Arg Ile Asn His Phe Glu Lys Ile Gln Ile Ile Pro Lys Ser
130 135 140

Ser Trp Ser Asp His Glu Ala Ser Ser Gly Val Ser Ser Ala Cys Pro
145 150 155 160

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

Lys Asn Ser Thr Tyr Pro Thr Ile Lys Lys Ser Tyr Asn Asn Thr Asn
180 185 190

Gln Glu Asp Leu Leu Val Leu Trp Gly Ile His His Pro Asn Asp Ala
195 200 205

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

Glu Ser Met Gly Val Tyr Gln Ile Leu Ala Ile Tyr Ser Thr Val Ala
530 535 540

Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe Trp Met
545 550 555 560

Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
565 570

<210> 83

<211> 575

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm dịch mã axit amin tổng hợp của trình tự mã hóa trong catxet biểu hiện sản phẩm thiết kế 737

<400> 83

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 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 Lys Phe Lys Tyr Pro Ala Leu Asn Val Thr Met
180 185 190

Pro Asn Asn Glu Lys Phe Asp Lys Leu Tyr Ile Trp Gly Val His His
195 200 205

Pro Gly Thr Asp Asn 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	Ile	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	
Ile	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	Ile	Gly	Thr	Tyr	Gln	Ile	Leu	Ser	Ile	Tyr	Ser	530	535	540	

Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly Leu Ser
545 550 555 560

Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
565 570 575

<210> 84

<211> 593

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm dịch mã axit amin tổng hợp của trình tự mã hóa trong catxet biểu hiện sản phẩm thiết kế 745

<400> 84

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 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 Glu 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 Tyr Gln Ile Leu Ser Ile
 545 550 555 560

Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly
565 570 575

Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys
580 585 590

Ile

<210> 85
<211> 405
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự SacI-Plasto 3(gốc)UTR tổng hợp

<400> 85
gagctctaag ttaaaatgct tcttcgtctc ctatttataa tatggtttgt tattgttaat 60
tttgttcttg tagaagagct taattaatcg ttgttggtat gaaatactat ttgtatgaga 120
tgaactgggtg taatgtaatt catttacata agtggagtca gaatcagaat gtttcctcca 180
taactaacta gacatgaaga cctgccgcgt acaattgtct tatatttgaa caactaaaat 240
tgaacatctt ttgccacaac tttataagtg gttaatatag ctcaaata tggtcaagtt 300
caatagatta ataatggaaa tatcagttat cgaaattcat taacaatcaa cttaacgtta 360
ttaactacta attttatatc atccccttg ataaatgata gtaca 405

<210> 86
<211> 1718
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> PDI sp - A/California/04/09 tổng hợp

<400> 86
atggcgaaaa acgttgcgat tttcggctta ttgttttctc ttcttgtgtt gggttcctctc 60
agatcttcgc tgacacatta tgtatagggt atcatgcgaa caattcaaca gacactgtag 120
acacagtact agaaaagaat gtaacagtaa cacactctgt taaccttcta gaagacaagc 180
ataacgggaa actatgcaaa ctaagagggg tagccccatt gcatttgggt aaatgtaaca 240
ttgctggctg gatcctggga aatccagagt gtgaatcact ctccacagca agctcatggt 300
cctacattgt ggaaacacct agttcagaca atggaacgtg ttaccagga gatttcacgc 360
attatgagga gctaagagag caattaagct cagtgtcatc atttgaaagg tttgagatat 420
tccccaaagac aagttcatgg cccaatcatg actcgaacaa aggtgtaacg gcagcatgtc 480
ctcatgctgg agcaaaaagc ttctacaaaa atttaatatg gctagttaaa aaaggaaatt 540
catacccaaaa gctcagcaaa tcctacatta atgataaagg gaaagaagtc ctcgtgctat 600

ggggcattca ccatccatct actagtgtctg accaacaag tctctatcag aatgcagata 660
 catatgtttt tgtgggggtca tcaagatata gcaagaagtt caagccggaa atagcaataa 720
 gacccaaagt gagggatcaa gaaggagaa tgaactatta ctggacacta gtagagccgg 780
 gagacaaaat aacattcgaa gcaactggaa atctagtggg accgagatat gcattcgcaa 840
 tggaaagaaa tgctggatct ggtattatca ttccagatac accagtccac gattgcaata 900
 caacttgtca aacaccaag ggtgctataa acaccagcct cccatttcag aatatacatc 960
 cgatcacaat tggaaaatgt ccaaaatatg taaaaagcac aaaattgaga ctggccacag 1020
 gattgaggaa tatcccgctt attcaatcta gaggactatt tggggccatt gccggtttca 1080
 ttgaaggggg gtggacaggg atggtagatg gatggtacgg ttatcaccat caaaatgagc 1140
 aggggtcagg atatgcagcc gacctgaaga gcacacagaa tgccattgac gagattacta 1200
 acaaagtaaa ttctgttatt gaaaagatga atacacagtt cacagcagta ggtaaagagt 1260
 tcaaccacct ggaaaaaaga atagagaatt taaataaaaa agttgatgat ggtttcctgg 1320
 acatttgac ttacaatgcc gaactgttgg ttctattgga aatgaaaga actttggact 1380
 accacgattc aaatgtgaag aacttatatg aaaaggtaag aagccagcta aaaaacaatg 1440
 ccaaggaaat tggaaacggc tgctttgaat ttaccacaa atgcgataac acgtgcatgg 1500
 aaagtgtcaa aaatgggact tatgactacc caaataactc agaggaagca aaattaaaca 1560
 gagaagaaat agatggggta aagctggaat caacaaggat ttaccagatt ttggcgatct 1620
 attcaactgt cgccagttca ttggtactgg tagtctccct gggggcaatc agtttctgga 1680
 tgtgctctaa tgggtctcta cagtgtagaa tatgtatt 1718

<210> 87

<211> 573

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> PDI sp - A/California 04/09 tổng hợp

<400> 87

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 Thr Leu Cys Ile Gly Tyr His
20 25 30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
35 40 45

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

Leu Cys Lys Leu Arg Gly Val Ala Pro Leu His Leu Gly Lys Cys Asn

23591

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

Asn Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala
 405 410 415
 Val Gly Lys Glu Phe Asn His Leu Glu Lys Arg Ile Glu Asn Leu Asn
 420 425 430
 Lys Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu
 435 440 445
 Leu Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Tyr His Asp Ser
 450 455 460
 Asn Val Lys Asn Leu Tyr Glu Lys Val Arg Ser Gln Leu Lys Asn Asn
 465 470 475 480
 Ala Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp
 485 490 495
 Asn Thr Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys
 500 505 510
 Tyr Ser Glu Glu Ala Lys Leu Asn Arg Glu Glu Ile Asp Gly Val Lys
 515 520 525
 Leu Glu Ser Thr Arg Ile Tyr Gln Ile Leu Ala Ile Tyr Ser Thr Val
 530 535 540
 Ala Ser Ser Leu Val Leu Val Val Ser Leu Gly Ala Ile Ser Phe Trp
 545 550 555 560
 Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
 565 570

<210> 88
 <211> 747
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Trình tự của promoter 2X35S tổng hợp

<400> 88
 gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60
 gaagacccaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120
 ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180
 tacaaatgcc atcattgcga taaaggaaag gccatcggtg aagatgcctc tgccgacagt 240
 ggtcccaaag atggaccccc acccagcagg agcatcggtg aaaaagaaga cggtccaacc 300
 acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac 360
 tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420
 agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480
 aagatagtggt aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540

atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc 600
 atcgtggaaa aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660
 tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata 720
 taaggaagtt catttcattt ggagagg 747

 <210> 89
 <211> 43
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> PacI-MCS-2X35S.c tổng hợp

 <400> 89
 aattgttaat taagtcgaca agcttgcatt cctgcaggtc aac 43

 <210> 90
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> CPMV 5(gốc)UTR-2X35S.r tổng hợp

 <400> 90
 tcaaaacctta ttaagatttt aatacctctc caaatgaaat gaacttcc 48

 <210> 91
 <211> 49
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> 2X35S-CPMV 5(gốc)UTR.c tổng hợp

 <400> 91
 ttggagaggt attaaaatct taatagggtt tgataaaagc gaacgtggg 49

 <210> 92
 <211> 44
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> ApaI-M prot.r tổng hợp

 <400> 92
 tctccatggg cccgacaaat ttgggcagaa tatacagaag ctta 44

 <210> 93
 <211> 3580
 <212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Sản phẩm thiết kế 747 tổng hợp, từ PacI đến AscI

 <400> 93

ttaattaagt	cgacaagctt	gcatgcctgc	aggtcaacat	ggtggagcac	gacacacttg	60
tctactccaa	aaatatcaaa	gatacagtct	cagaagacca	aagggcaatt	gagacttttc	120
aacaaagggg	aatatccgga	aacctcctcg	gattccattg	cccagctatc	tgtcacttta	180
ttgtgaagat	agtggaaaag	gaagggtggc	cctacaaatg	ccatcattgc	gataaaggaa	240
aggccatcgt	tgaagatgcc	tctgccgaca	gtgggtccaa	agatggaccc	ccacccacga	300
ggagcatcgt	ggaaaaagaa	gacgttccaa	ccacgtcttc	aaagcaagtg	gattgatgtg	360
ataacatggg	ggagcacgac	acacttgtct	actccaaaaa	tatcaaagat	acagtctcag	420
aagaccaaag	ggcaattgag	acttttcaac	aaagggtaat	atccggaaac	ctcctcggat	480
tccattgccc	agctatctgt	cactttattg	tgaagatagt	ggaaaaggaa	ggtggctcct	540
acaaatgcca	tcattgcat	aaaggaaagg	ccatcgttga	agatgcctct	gccgacagtg	600
gtcccaaaga	tggaccccca	cccacgagga	gcatcgtgga	aaaagaagac	gttccaacca	660
cgtcttcaaa	gcaagtggat	tgatgtgata	tctccactga	cgtaagggat	gacgcacaat	720
cccactatcc	ttcgcaagac	ccttcctcta	tataaggaag	ttcatttcat	ttggagaggt	780
attaaaatct	taataggttt	tgataaaagc	gaacgtgggg	aaacccgaac	caaaccttct	840
tctaaactct	ctctcatctc	tcttaaagca	aacttctctc	ttgtctttct	tgcgtgagcg	900
atcttcaacg	ttgtcagatc	gtgcttcggc	accagtacaa	cgttttcttt	cactgaagcg	960
aatcaaaga	tctctttgtg	gacacgtagt	gcggcgccat	taaataacgt	gtacttgtcc	1020
tattcttgtc	ggtgtggtct	tgggaaaaga	aagcttgctg	gaggctgctg	ttcagcccca	1080
tacattactt	gttacgattc	tgctgacttt	cggcgggtgc	aatatctcta	cttctgcttg	1140
acgaggtatt	gttgccgtga	cttctttctt	cttcttcttg	ctgattggtt	ctataagaaa	1200
tctagtatth	tctttgaaac	agagttttcc	cgtgggtttc	gaacttggag	aaagattggt	1260
aagcttctgt	atattctgcc	caaatttgtc	gggcccatgg	cgaaaaacgt	tgcgattttc	1320
ggcttattgt	tttctcttct	tgtgttggtt	ccttctcaga	tcttcgctga	tcgaatctgc	1380
actggaataa	catcttcaaa	ctcacctcat	gtgggtcaaaa	cagccactca	aggggaggtc	1440
aatgtgactg	gtgtgatacc	actaacaaca	acaccaacaa	aatcttattt	tgcaaatctc	1500
aaaggaacaa	ggaccagagg	gaaactatgc	ccagactgtc	tcaactgcac	agatctggat	1560
gtggcttttg	gcagaccaat	gtgtgtgggg	accacacctt	cggcgaaggc	ttcaatactc	1620
cacgaagtca	aacctgttac	atccgggtgc	tttctataaa	tgcacgacag	aacaaaaatc	1680
aggcaactac	ccaatcttct	cagaggatat	gaaaatatca	ggctatcaac	ccaaaacgtc	1740
atcgatgcgg	aaaaggcacc	aggaggaccc	tacagacttg	gaacctcagg	atcttgccct	1800
aacgctacca	gtaagagcgg	atttttcgca	acaatggctt	gggctgtccc	aaaggacaac	1860

aacaaaaatg caacgaaccc actaacagta gaagtacat acatttgtac agaaggggaa 1920
gaccaaataca ctgtttgggg gttccattca gataacaaaa cccaaatgaa gaacctctat 1980
ggagactcaa atcctcaaaa gttcacctca tctgctaata gagtaaccac acactatgtt 2040
tctcagattg gcagcttccc agatcaaaca gaagacggag gactaccaca aagcggcagg 2100
attgtttgtt attacatgat gcaaaaacct gggaaaacag gaacaattgt ctaccaaaaga 2160
ggtgttttgt tgcctcaaaa ggtgtggtgc gcgagtggca ggagcaaagt aataaaaggg 2220
tccttgcctt taattggtga agcagattgc cttcatgaaa aatacggtag attaaacaaa 2280
agcaagcctt actacacagg agaacatgca aaagccatag gaaattgccc aatatgggtg 2340
aaaacacctt tgaagctcgc caatggaacc aaatatagac ctcttgcaaa actattaaag 2400
gaaaggggtt tcttcggagc tattgctggt ttctagaag gaggatggga aggaatgatt 2460
gcaggctggc acggatacac atctcacgga gcacatggag tggcagtggc ggcggacctt 2520
aagagtacgc aagaagctat aaacaagata aaaaaaatc tcaattcttt gagtgaagcta 2580
gaagtaaaga atcttcaaag actaagtggg gccatggatg aactccacaa cgaaatactc 2640
gagctggatg agaaagtgga tgatctcaga gctgacacta taagctcgca aatagaactt 2700
gcagtcttgc tttccaacga aggaataata aacagtgaag atgagcatct attggcactt 2760
gagagaaaac taaagaaaat gctgggtccc tctgctgtag agataggaaa tggatgcttc 2820
gaaaccaaac acaagtgcaa ccagacctgc ttagacagga tagctgctgg cacctttaat 2880
gcaggagaat tttctctccc cacttttgat tcaactgaaca ttactgctgc atctttaaat 2940
gatgatggat tggataacta ccaaatactg tcaatttatt caacagtggc gagttcccta 3000
gcactggcaa tcatgatggc tgggtctatct ttatggatgt gctccaatgg atcggtacaa 3060
tgcagaatth gcatttaaag gcctatthtc tttagtttga atttactgtt attcgggtgtg 3120
catttctatg tttggtgagc ggttttctgt gctcagagtg tgtttatttt atgtaattta 3180
atttctttgt gagctcctgt ttagcaggtc gtcccttcag caaggacaca aaaagatttt 3240
aattttatta aaaaaaaaaa aaaaaaagac cggaatttcg atatcaagct tatcgacctg 3300
cagatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt gccgggtctt 3360
cgatgattat catataatth ctgttgaatt acgttaagca tgtaataatt aacatgtaat 3420
gcatgacgtt atttatgaga tgggttttth tgattagagt cccgcaatta tacatttaat 3480
acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcg gcgggtgtcat 3540
ctatgttact agattctaga gtctcaagct tcggcgcgcc 3580

<210> 94

<211> 569

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Virut cúm B B/Florida/4/2006 tổng hợp

<400> 94

```

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 Tyr Phe Ala Asn Leu Lys Gly Thr Arg
          35          40          45

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

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

Ala Ser Ile Leu His Glu Val Lys 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 Asn Ile Arg Leu Ser Thr Gln Asn Val Ile Asp Ala Glu
          115          120          125

Lys Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser Gly Ser Cys Pro
          130          135          140

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

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

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

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

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

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

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

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

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

```

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

<210> 95
 <211> 570
 <212> PRT
 <213> Trình tự nhân tạo

<220>

<223> Virut cúm B B/Malaysia/2506/2004 tổng hợp

<400> 95

```

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

Pro Lys Asn Asp Asn Asn Lys Thr Ala Thr Asn Ser 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> 96
 <211> 548
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Virut cúm A A/Brisbane/59/2007(H1N1)

<400> 96

Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr	Val	1	5	10	15
Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn	Leu	20	25	30	
Leu	Glu	Asn	Ser	His	Asn	Gly	Lys	Leu	Cys	Leu	Leu	Lys	Gly	Ile	Ala	35	40	45	
Pro	Leu	Gln	Leu	Gly	Asn	Cys	Ser	Val	Ala	Gly	Trp	Ile	Leu	Gly	Asn	50	55	60	
Pro	Glu	Cys	Glu	Leu	Leu	Ile	Ser	Lys	Glu	Ser	Trp	Ser	Tyr	Ile	Val	65	70	75	80
Glu	Lys	Pro	Asn	Pro	Glu	Asn	Gly	Thr	Cys	Tyr	Pro	Gly	His	Phe	Ala	85	90	95	
Asp	Tyr	Glu	Glu	Leu	Arg	Glu	Gln	Leu	Ser	Ser	Val	Ser	Ser	Phe	Glu	100	105	110	
Arg	Phe	Glu	Ile	Phe	Pro	Lys	Glu	Ser	Ser	Trp	Pro	Asn	His	Thr	Val	115	120	125	
Thr	Gly	Val	Ser	Ala	Ser	Cys	Ser	His	Asn	Gly	Glu	Ser	Ser	Phe	Tyr	130	135	140	
Arg	Asn	Leu	Leu	Trp	Leu	Thr	Gly	Lys	Asn	Gly	Leu	Tyr	Pro	Asn	Leu	145	150	155	160
Ser	Lys	Ser	Tyr	Ala	Asn	Asn	Lys	Glu	Lys	Glu	Val	Leu	Val	Leu	Trp	165	170	175	
Gly	Val	His	His	Pro	Pro	Asn	Ile	Gly	Asn	Gln	Lys	Ala	Leu	Tyr	His	180	185	190	
Thr	Glu	Asn	Ala	Tyr	Val	Ser	Val	Val	Ser	Ser	His	Tyr	Ser	Arg	Lys	195	200	205	
Phe	Thr	Pro	Glu	Ile	Ala	Lys	Arg	Pro	Lys	Val	Arg	Asp	Gln	Glu	Gly	210	215	220	
Arg	Ile	Asn	Tyr	Tyr	Trp	Thr	Leu	Leu	Glu	Pro	Gly	Asp	Thr	Ile	Ile	225	230	235	240
Phe	Glu	Ala	Asn	Gly	Asn	Leu	Ile	Ala	Pro	Arg	Tyr	Ala	Phe	Ala	Leu	245	250	255	
Ser	Arg	Gly	Phe	Gly	Ser	Gly	Ile	Ile	Asn	Ser	Asn	Ala	Pro	Met	Asp	260	265	270	
Lys	Cys	Asp	Ala	Lys	Cys	Gln	Thr	Pro	Gln	Gly	Ala	Ile	Asn	Ser	Ser	275	280	285	
Leu	Pro	Phe	Gln	Asn	Val	His	Pro	Val	Thr	Ile	Gly	Glu	Cys	Pro	Lys	290	295	300	
Tyr	Val	Arg	Ser	Ala	Lys	Leu	Arg	Met	Val	Thr	Gly	Leu	Arg	Asn	Ile	305	310	315	320
Pro	Ser	Ile	Gln	Ser	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile				

				325						330					335		
Glu	Gly	Gly	Trp	Thr	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His		
			340					345					350				
Gln	Asn	Glu	Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Gln	Lys	Ser	Thr	Gln		
		355					360					365					
Asn	Ala	Ile	Asn	Gly	Ile	Thr	Asn	Lys	Val	Asn	Ser	Val	Ile	Glu	Lys		
	370					375					380						
Met	Asn	Thr	Gln	Phe	Thr	Ala	Val	Gly	Lys	Glu	Phe	Asn	Lys	Leu	Glu		
385					390					395					400		
Arg	Arg	Met	Glu	Asn	Leu	Asn	Lys	Lys	Val	Asp	Asp	Gly	Phe	Ile	Asp		
				405					410					415			
Ile	Trp	Thr	Tyr	Asn	Ala	Glu	Leu	Leu	Val	Leu	Leu	Glu	Asn	Glu	Arg		
			420					425					430				
Thr	Leu	Asp	Phe	His	Asp	Ser	Asn	Val	Lys	Asn	Leu	Tyr	Glu	Lys	Val		
		435					440					445					
Lys	Ser	Gln	Leu	Lys	Asn	Asn	Ala	Lys	Glu	Ile	Gly	Asn	Gly	Cys	Phe		
	450					455					460						
Glu	Phe	Tyr	His	Lys	Cys	Asn	Asp	Glu	Cys	Met	Glu	Ser	Val	Lys	Asn		
465					470					475					480		
Gly	Thr	Tyr	Asp	Tyr	Pro	Lys	Tyr	Ser	Glu	Glu	Ser	Lys	Leu	Asn	Arg		
				485					490					495			
Glu	Lys	Ile	Asp	Gly	Val	Lys	Leu	Glu	Ser	Met	Gly	Val	Tyr	Gln	Ile		
			500					505					510				
Leu	Ala	Ile	Tyr	Ser	Thr	Val	Ala	Ser	Ser	Leu	Val	Leu	Leu	Val	Ser		
		515					520					525					
Leu	Gly	Ala	Ile	Ser	Phe	Trp	Met	Cys	Ser	Asn	Gly	Ser	Leu	Gln	Cys		
	530					535					540						

Arg Ile Cys Ile
545

<210> 97
 <211> 548
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Virut cúm A A/Solomon Islands/3/2006(H1N1) tổng hợp

<400> 97

Asp	Thr	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Asp	Thr	Val
1				5					10					15	
Asp	Thr	Val	Leu	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ser	Val	Asn	Leu
			20					25					30		
Leu	Glu	Asp	Ser	His	Asn	Gly	Lys	Leu	Cys	Leu	Leu	Lys	Gly	Ile	Ala

23591

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

Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys
 370 375 380
 Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu Glu
 385 390 395 400
 Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Ile Asp
 405 410 415
 Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg
 420 425 430
 Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
 435 440 445
 Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
 450 455 460
 Glu Phe Tyr His Lys Cys Asn Asp Glu Cys Met Glu Ser Val Lys Asn
 465 470 475 480
 Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
 485 490 495
 Glu Lys Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr Gln Ile
 500 505 510
 Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser
 515 520 525
 Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys
 530 535 540

Arg Ile Cys Ile
 545

<210> 98
 <211> 548
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Virut cúm A A/New Caledonia/20/1999(H1N1) tổng hợp

<400> 98

Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val
 1 5 10 15
 Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn Leu
 20 25 30
 Leu Glu Asp Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala
 35 40 45
 Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn
 50 55 60
 Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile Val
 65 70 75 80

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

Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu Asp
 405 410 415
 Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu Arg
 420 425 430
 Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys Val
 435 440 445
 Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys Phe
 450 455 460
 Glu Phe Tyr His Lys Cys Asn Asn Glu Cys Met Glu Ser Val Lys Asn
 465 470 475 480
 Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn Arg
 485 490 495
 Glu Lys Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr Gln Ile
 500 505 510
 Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val Ser
 515 520 525
 Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys
 530 535 540
 Arg Ile Cys Ile
 545
 <210> 99
 <211> 547
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> Virut cúm A A/Singapore/1/1957(H2N2)tổng hợp
 <400> 99
 Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Lys Val
 1 5 10 15
 Asp Thr Ile Leu Glu Arg Asn Val Thr Val Thr His Ala Lys Asp Ile
 20 25 30
 Leu Glu Lys Thr His Asn Gly Lys Leu Cys Lys Leu Asn Gly Ile Pro
 35 40 45
 Pro Leu Glu Leu Gly Asp Cys Ser Ile Ala Gly Trp Leu Leu Gly Asn
 50 55 60
 Pro Glu Cys Asp Arg Leu Leu Ser Val Pro Glu Trp Ser Tyr Ile Met
 65 70 75 80
 Glu Lys Glu Asn Pro Arg Asp Gly Leu Cys Tyr Pro Gly Ser Phe Asn
 85 90 95
 Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Ser Val Lys His Phe Glu
 100 105 110

Lys Val Lys Ile Leu Pro Lys Asp Arg Trp Thr Gln His Thr Thr Thr
 115 120 125
 Gly Gly Ser Arg Ala Cys Ala Val Ser Gly Asn Pro Ser Phe Phe Arg
 130 135 140
 Asn Met Val Trp Leu Thr Lys Lys Glu Ser Asn Tyr Pro Val Ala Lys
 145 150 155 160
 Gly Ser Tyr Asn Asn Thr Ser Gly Glu Gln Met Leu Ile Ile Trp Gly
 165 170 175
 Val His His Pro Asn Asp Glu Thr Glu Gln Arg Thr Leu Tyr Gln Asn
 180 185 190
 Val Gly Thr Tyr Val Ser Val Gly Thr Ser Thr Leu Asn Lys Arg Ser
 195 200 205
 Thr Pro Asp Ile Ala Thr Arg Pro Lys Val Asn Gly Leu Gly Ser Arg
 210 215 220
 Met Glu Phe Ser Trp Thr Leu Leu Asp Met Trp Asp Thr Ile Asn Phe
 225 230 235 240
 Glu Ser Thr Gly Asn Leu Ile Ala Pro Glu Tyr Gly Phe Lys Ile Ser
 245 250 255
 Lys Arg Gly Ser Ser Gly Ile Met Lys Thr Glu Gly Thr Leu Glu Asn
 260 265 270
 Cys Glu Thr Lys Cys Gln Thr Pro Leu Gly Ala Ile Asn Thr Thr Leu
 275 280 285
 Pro Phe His Asn Val His Pro Leu Thr Ile Gly Glu Cys Pro Lys Tyr
 290 295 300
 Val Lys Ser Glu Lys Leu Val Leu Ala Thr Gly Leu Arg Asn Val Pro
 305 310 315 320
 Gln Ile Glu Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe Ile Glu
 325 330 335
 Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly Tyr His His Ser
 340 345 350
 Asn Asp Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu Ser Thr Gln Lys
 355 360 365
 Ala Phe Asp Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu Lys Met
 370 375 380
 Asn Thr Gln Phe Glu Ala Val Gly Lys Glu Phe Ser Asn Leu Glu Arg
 385 390 395 400
 Arg Leu Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp Val
 405 410 415
 Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg Thr
 420 425 430
 Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val Arg

435 440 445
 Met Gln Leu Arg Asp Asn Val Lys Glu Leu Gly Asn Gly Cys Phe Glu
 450 455 460
 Phe Tyr His Lys Cys Asp Asp Glu Cys Met Asn Ser Val Lys Asn Gly
 465 470 475 480
 Thr Tyr Asp Tyr Pro Lys Tyr Glu Glu Glu Ser Lys Leu Asn Arg Asn
 485 490 495
 Glu Ile Lys Gly Val Lys Leu Ser Ser Met Gly Val Tyr Gln Ile Leu
 500 505 510
 Ala Ile Tyr Ala Thr Val Ala Gly Ser Leu Ser Leu Ala Ile Met Met
 515 520 525
 Ala Gly Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg
 530 535 540
 Ile Cys Ile
 545
 <210> 100
 <211> 541
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> [Virut cúm A A/Brisbane/10/2007(H3N2) tổng hợp
 <400> 100
 Thr Ala Thr Leu Cys Leu Gly His His Ala Val Pro Asn Gly Thr Ile
 1 5 10 15
 Val Lys Thr Ile Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr Glu
 20 25 30
 Leu Val Gln Ser Ser Ser Thr Gly Glu Ile Cys Asp Ser Pro His Gln
 35 40 45
 Ile Leu Asp Gly Glu Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly Asp
 50 55 60
 Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val Glu
 65 70 75 80
 Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp Tyr
 85 90 95
 Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe Asn
 100 105 110
 Asn Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser Ser
 115 120 125
 Ala Cys Ile Arg Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn Trp
 130 135 140
 Leu Thr His Leu Lys Phe Lys Tyr Pro Ala Leu Asn Val Thr Met Pro

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

Val Tyr Arg Asp Glu Ala Leu Asn Asn Arg Phe Gln Ile Lys Gly Val
485 490 495

Glu Leu Lys Ser Gly Tyr Lys Asp Trp Ile Leu Trp Ile Ser Phe Ala
500 505 510

Ile Ser Cys Phe Leu Leu Cys Val Ala Leu Leu Gly Phe Ile Met Trp
515 520 525

Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
530 535 540

<210> 101

<211> 541

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> A/Wisconsin/67e5/2005(H3)tổng hợp

<400> 101

Thr Ala Thr Leu Cys Leu Gly His His Ala Val Pro Asn Gly Thr Ile
1 5 10 15

Val Lys Thr Ile Thr Asn Asp Gln Ile Glu Val Thr Asn Ala Thr Glu
20 25 30

Leu Val Gln Ser Ser Ser Thr Gly Gly Ile Cys Asp Ser Pro His Gln
35 40 45

Ile Leu Asp Gly Glu Asn Cys Thr Leu Ile Asp Ala Leu Leu Gly Asp
50 55 60

Pro Gln Cys Asp Gly Phe Gln Asn Lys Lys Trp Asp Leu Phe Val Glu
65 70 75 80

Arg Ser Lys Ala Tyr Ser Asn Cys Tyr Pro Tyr Asp Val Pro Asp Tyr
85 90 95

Ala Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe Asn
100 105 110

Asp Glu Ser Phe Asn Trp Thr Gly Val Thr Gln Asn Gly Thr Ser Ser
115 120 125

Ala Cys Lys Arg Arg Ser Asn Asn Ser Phe Phe Ser Arg Leu Asn Trp
130 135 140

Leu Thr His Leu Lys Phe Lys Tyr Pro Ala Leu Asn Val Thr Met Pro
145 150 155 160

Asn Asn Glu Lys Phe Asp Lys Leu Tyr Ile Trp Gly Val His His Pro
165 170 175

Gly Thr Asp Asn Asp Gln Ile Phe Leu His Ala Gln Ala Ser Gly Arg
180 185 190

Ile Thr Val Ser Thr Lys Arg Ser Gln Gln Thr Val Ile Pro Asn Ile
195 200 205

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

Ala Cys Gln Lys Gly Asn Ile Arg Cys Asn Ile Cys Ile
530 535 540

<210> 102

<211> 551

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Virut cúm A A/Anhui/1/2005 (H5N1) tổng hợp

<400> 102

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
1 5 10 15

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
20 25 30

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
35 40 45

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
50 55 60

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
65 70 75 80

Glu Lys Ala Asn Pro Ala Asn Asp Leu Cys Tyr Pro Gly Asn Phe Asn
85 90 95

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
100 105 110

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
115 120 125

Ser Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Thr Pro Ser Phe Phe
130 135 140

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Asn Thr Tyr Pro Thr Ile
145 150 155 160

Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Ile Leu Trp
165 170 175

Gly Ile His His Ser Asn Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln
180 185 190

Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg
195 200 205

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

Arg Met Asp Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
225 230 235 240

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
245 250 255

23591

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

<210> 103

<211> 552

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> [Virut cúm A A/Vietnam/1194/2004 (H5N1) tổng hợp

<400> 103

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
 1 5 10 15
 Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
 20 25 30
 Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
 35 40 45
 Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
 50 55 60
 Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
 65 70 75 80
 Glu Lys Ala Asn Pro Val Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn
 85 90 95
 Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
 100 105 110
 Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Ser His Glu Ala Ser
 115 120 125
 Leu Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe
 130 135 140
 Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
 145 150 155 160
 Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
 165 170 175
 Gly Ile His His Pro Lys Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln
 180 185 190
 Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg
 195 200 205
 Leu Val Pro Arg Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
 210 215 220
 Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
 225 230 235 240
 Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
 245 250 255
 Val Lys Lys Gly Asp Ser Thr Ile Met Lys Ser Glu Leu Glu Tyr Gly
 260 265 270
 Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
 275 280 285
 Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys

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

<210> 104

<211> 553

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> A/Indonesia/5/2005(H5N1) tổng hợp

<400> 104

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val

1		5		10		15													
Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile				
			20					25					30						
Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Asp	Leu	Asp	Gly	Val	Lys				
		35					40					45							
Pro	Leu	Ile	Leu	Arg	Asp	Cys	Ser	Val	Ala	Gly	Trp	Leu	Leu	Gly	Asn				
	50					55					60								
Pro	Met	Cys	Asp	Glu	Phe	Ile	Asn	Val	Pro	Glu	Trp	Ser	Tyr	Ile	Val				
65					70					75					80				
Glu	Lys	Ala	Asn	Pro	Thr	Asn	Asp	Leu	Cys	Tyr	Pro	Gly	Ser	Phe	Asn				
				85					90					95					
Asp	Tyr	Glu	Glu	Leu	Lys	His	Leu	Leu	Ser	Arg	Ile	Asn	His	Phe	Glu				
			100					105					110						
Lys	Ile	Gln	Ile	Ile	Pro	Lys	Ser	Ser	Trp	Ser	Asp	His	Glu	Ala	Ser				
		115					120					125							
Ser	Gly	Val	Ser	Ser	Ala	Cys	Pro	Tyr	Leu	Gly	Ser	Pro	Ser	Phe	Phe				
	130					135					140								
Arg	Asn	Val	Val	Trp	Leu	Ile	Lys	Lys	Asn	Ser	Thr	Tyr	Pro	Thr	Ile				
145					150					155					160				
Lys	Lys	Ser	Tyr	Asn	Asn	Thr	Asn	Gln	Glu	Asp	Leu	Leu	Val	Leu	Trp				
				165					170					175					
Gly	Ile	His	His	Pro	Asn	Asp	Ala	Ala	Glu	Gln	Thr	Arg	Leu	Tyr	Gln				
			180					185					190						
Asn	Pro	Thr	Thr	Tyr	Ile	Ser	Ile	Gly	Thr	Ser	Thr	Leu	Asn	Gln	Arg				
		195					200					205							
Leu	Val	Pro	Lys	Ile	Ala	Thr	Arg	Ser	Lys	Val	Asn	Gly	Gln	Ser	Gly				
	210					215					220								
Arg	Met	Glu	Phe	Phe	Trp	Thr	Ile	Leu	Lys	Pro	Asn	Asp	Ala	Ile	Asn				
225					230					235					240				
Phe	Glu	Ser	Asn	Gly	Asn	Phe	Ile	Ala	Pro	Glu	Tyr	Ala	Tyr	Lys	Ile				
				245					250					255					
Val	Lys	Lys	Gly	Asp	Ser	Ala	Ile	Met	Lys	Ser	Glu	Leu	Glu	Tyr	Gly				
			260					265					270						
Asn	Cys	Asn	Thr	Lys	Cys	Gln	Thr	Pro	Met	Gly	Ala	Ile	Asn	Ser	Ser				
		275					280					285							
Met	Pro	Phe	His	Asn	Ile	His	Pro	Leu	Thr	Ile	Gly	Glu	Cys	Pro	Lys				
	290					295					300								
Tyr	Val	Lys	Ser	Asn	Arg	Leu	Val	Leu	Ala	Thr	Gly	Leu	Arg	Asn	Ser				
305					310					315					320				
Pro	Gln	Arg	Glu	Ser	Arg	Arg	Lys	Lys	Arg	Gly	Leu	Phe	Gly	Ala	Ile				
				325					330					335					

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

Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
355 360 365

Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
370 375 380

Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
385 390 395 400

Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
405 410 415

Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
420 425 430

Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
435 440 445

Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
450 455 460

Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
465 470 475 480

Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala
485 490 495

Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
500 505 510

Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
515 520 525

Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
530 535 540

Ser Leu Gln Cys Arg Ile Cys Ile Lys
545 550

<210> 105

<211> 548

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế 900

<400> 105

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
1 5 10 15

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
20 25 30

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
35 40 45

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

Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile Ile Asp Lys
 370 375 380
 Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn Asn Leu Glu
 385 390 395 400
 Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly Phe Leu Asp
 405 410 415
 Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu Asn Glu Arg
 420 425 430
 Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val
 435 440 445
 Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe
 450 455 460
 Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser Ile Arg Asn
 465 470 475 480
 Gly Thr Tyr Asn Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys Arg
 485 490 495
 Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln Ile
 500 505 510
 Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met
 515 520 525
 Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys
 530 535 540
 Arg Ile Cys Ile
 545
 <210> 106
 <211> 565
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> Sản phẩm thiết kế 745
 <400> 106
 Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
 1 5 10 15
 Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
 20 25 30
 Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
 35 40 45
 Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
 50 55 60
 Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
 65 70 75 80

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

				405					410					415					
Glu	Arg	Arg	Ile	Glu	Asn	Leu	Asn	Lys	Lys	Met	Glu	Asp	Gly	Phe	Leu				
			420					425					430						
Asp	Val	Trp	Thr	Tyr	Asn	Ala	Glu	Leu	Leu	Val	Leu	Met	Glu	Asn	Glu				
		435					440					445							
Arg	Thr	Leu	Asp	Phe	His	Asp	Ser	Asn	Val	Lys	Asn	Leu	Tyr	Asp	Lys				
	450					455					460								
Val	Arg	Leu	Gln	Leu	Arg	Asp	Asn	Ala	Lys	Glu	Leu	Gly	Asn	Gly	Cys				
465					470					475					480				
Phe	Glu	Phe	Tyr	His	Lys	Cys	Asp	Asn	Glu	Cys	Met	Glu	Ser	Ile	Arg				
				485					490					495					
Asn	Gly	Thr	Tyr	Asn	Tyr	Pro	Gln	Tyr	Ser	Glu	Glu	Ala	Arg	Leu	Lys				
			500					505					510						
Arg	Glu	Glu	Ile	Ser	Gly	Val	Lys	Leu	Glu	Ser	Ile	Gly	Thr	Tyr	Gln				
		515					520					525							
Ile	Leu	Ser	Ile	Tyr	Ser	Thr	Val	Ala	Ser	Ser	Leu	Ala	Leu	Ala	Ile				
	530					535					540								
Met	Met	Ala	Gly	Leu	Ser	Leu	Trp	Met	Cys	Ser	Asn	Gly	Ser	Leu	Gln				
545					550					555					560				
Cys	Arg	Ile	Cys	Ile															
				565															

<210> 107

<211> 545

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế 910

<400> 107

Asp	Gln	Ile	Cys	Ile	Gly	Tyr	His	Ala	Asn	Asn	Ser	Thr	Glu	Gln	Val				
1				5					10					15					
Asp	Thr	Ile	Met	Glu	Lys	Asn	Val	Thr	Val	Thr	His	Ala	Gln	Asp	Ile				
			20					25					30						
Leu	Glu	Lys	Thr	His	Asn	Gly	Lys	Leu	Cys	Asp	Ser	Pro	His	Gln	Ile				
		35				40						45							
Leu	Asp	Gly	Glu	Asn	Cys	Thr	Leu	Ile	Asp	Ala	Leu	Leu	Gly	Asp	Pro				
	50					55					60								
Gln	Cys	Asp	Gly	Phe	Gln	Asn	Lys	Lys	Trp	Asp	Leu	Phe	Val	Glu	Arg				
65					70					75					80				
Ser	Lys	Ala	Tyr	Ser	Asn	Cys	Tyr	Pro	Tyr	Asp	Val	Pro	Asp	Tyr	Ala				
				85					90					95					
Ser	Leu	Arg	Ser	Leu	Val	Ala	Ser	Ser	Gly	Thr	Leu	Glu	Phe	Asn	Asn				

100					105					110					
Glu	Ser	Phe	Asn	Trp	Thr	Gly	Val	Thr	Gln	Asn	Gly	Thr	Ser	Ser	Ala
		115					120					125			
Cys	Ile	Arg	Arg	Ser	Asn	Asn	Ser	Phe	Phe	Ser	Arg	Leu	Asn	Trp	Leu
	130					135					140				
Thr	His	Leu	Lys	Phe	Lys	Tyr	Pro	Ala	Leu	Asn	Val	Thr	Met	Pro	Asn
145					150					155					160
Asn	Glu	Lys	Phe	Asp	Lys	Leu	Tyr	Ile	Trp	Gly	Val	His	His	Pro	Gly
				165					170					175	
Thr	Asp	Asn	Asp	Gln	Ile	Phe	Leu	Tyr	Ala	Gln	Ala	Ser	Gly	Arg	Ile
			180					185					190		
Thr	Val	Ser	Thr	Lys	Arg	Ser	Gln	Gln	Thr	Val	Ile	Pro	Asn	Ile	Gly
	195						200					205			
Ser	Arg	Pro	Arg	Val	Arg	Asn	Ile	Pro	Ser	Arg	Ile	Ser	Ile	Tyr	Trp
	210					215					220				
Thr	Ile	Val	Lys	Pro	Gly	Asp	Ile	Leu	Leu	Ile	Asn	Ser	Thr	Gly	Asn
225					230					235					240
Leu	Ile	Ala	Pro	Arg	Gly	Tyr	Phe	Lys	Ile	Arg	Ser	Gly	Lys	Ser	Ser
				245					250					255	
Ile	Met	Arg	Ser	Asp	Ala	Pro	Ile	Gly	Lys	Cys	Asn	Thr	Lys	Cys	Gln
			260					265					270		
Thr	Pro	Met	Gly	Ala	Ile	Asn	Ser	Ser	Met	Pro	Phe	His	Asn	Ile	His
		275					280					285			
Pro	Leu	Thr	Ile	Gly	Glu	Cys	Pro	Lys	Tyr	Val	Lys	Ser	Asn	Arg	Leu
	290					295					300				
Val	Leu	Ala	Thr	Gly	Leu	Arg	Asn	Ser	Pro	Gln	Arg	Glu	Ser	Arg	Arg
305					310					315					320
Lys	Lys	Arg	Gly	Leu	Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Gly	Gly
			325						330					335	
Trp	Gln	Gly	Met	Val	Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Ser	Asn	Glu
			340					345					350		
Gln	Gly	Ser	Gly	Tyr	Ala	Ala	Asp	Lys	Glu	Ser	Thr	Gln	Lys	Ala	Ile
		355					360					365			
Asp	Gly	Val	Thr	Asn	Lys	Val	Asn	Ser	Ile	Ile	Asp	Lys	Met	Asn	Thr
	370					375					380				
Gln	Phe	Glu	Ala	Val	Gly	Arg	Glu	Phe	Asn	Asn	Leu	Glu	Arg	Arg	Ile
385					390					395					400
Glu	Asn	Leu	Asn	Lys	Lys	Met	Glu	Asp	Gly	Phe	Leu	Asp	Val	Trp	Thr
				405					410					415	
Tyr	Asn	Ala	Glu	Leu	Leu	Val	Leu	Met	Glu	Asn	Glu	Arg	Thr	Leu	Asp
			420					425					430		

Phe His Asp Ser Asn Val Lys Asn Leu Tyr Asp Lys Val Arg Leu Gln
 435 440 445
 Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn Gly Cys Phe Glu Phe Tyr
 450 455 460
 His Lys Cys Asp Asn Glu Cys Met Glu Ser Ile Arg Asn Gly Thr Tyr
 465 470 475 480
 Asn Tyr Pro Gln Tyr Ser Glu Glu Ala Arg Leu Lys Arg Glu Glu Ile
 485 490 495
 Ser Gly Val Lys Leu Glu Ser Ile Gly Thr Tyr Gln Ile Leu Ser Ile
 500 505 510
 Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly
 515 520 525
 Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys
 530 535 540

Ile
545

<210> 108
 <211> 556
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Sản phẩm thiết kế 920

<400> 108

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
 1 5 10 15
 Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
 20 25 30
 Leu Glu Lys Thr His Asn Gly Lys Leu Cys Pro Asp Cys Leu Asn Cys
 35 40 45
 Thr Asp Leu Asp Val Ala Leu Gly Arg Pro Met Cys Val Gly Thr Thr
 50 55 60
 Pro Ser Ala Lys Ala Ser Ile Leu His Glu Val Lys Pro Val Thr Ser
 65 70 75 80
 Gly Cys Phe Pro Ile Met His Asp Arg Thr Lys Ile Arg Gln Leu Pro
 85 90 95
 Asn Leu Leu Arg Gly Tyr Glu Asn Ile Arg Leu Ser Thr Gln Asn Val
 100 105 110
 Ile Asp Ala Glu Lys Ala Pro Gly Gly Pro Tyr Arg Leu Gly Thr Ser
 115 120 125
 Gly Ser Cys Pro Asn Ala Thr Ser Lys Ser Gly Phe Phe Ala Thr Met
 130 135 140

Ala	Trp	Ala	Val	Pro	Lys	Asp	Asn	Asn	Lys	Asn	Ala	Thr	Asn	Pro	Leu	145	150	155	160
Thr	Val	Glu	Val	Pro	Tyr	Ile	Cys	Thr	Glu	Gly	Glu	Asp	Gln	Ile	Thr	165	170	175	
Val	Trp	Gly	Phe	His	Ser	Asp	Asn	Lys	Thr	Gln	Met	Lys	Asn	Leu	Tyr	180	185	190	
Gly	Asp	Ser	Asn	Pro	Gln	Lys	Phe	Thr	Ser	Ser	Ala	Asn	Gly	Val	Thr	195	200	205	
Thr	His	Tyr	Val	Ser	Gln	Ile	Gly	Ser	Phe	Pro	Asp	Gln	Thr	Glu	Asp	210	215	220	
Gly	Gly	Leu	Pro	Gln	Ser	Gly	Arg	Ile	Val	Val	Asp	Tyr	Met	Met	Gln	225	230	235	240
Lys	Pro	Gly	Lys	Thr	Gly	Thr	Ile	Val	Tyr	Gln	Arg	Gly	Val	Leu	Leu	245	250	255	
Pro	Gln	Lys	Val	Trp	Cys	Ala	Ser	Gly	Arg	Ser	Lys	Val	Ile	Lys	Gly	260	265	270	
Ser	Leu	Pro	Leu	Ile	Gly	Glu	Ala	Asp	Cys	Gln	Thr	Pro	Met	Gly	Ala	275	280	285	
Ile	Asn	Ser	Ser	Met	Pro	Phe	His	Asn	Ile	His	Pro	Leu	Thr	Ile	Gly	290	295	300	
Glu	Cys	Pro	Lys	Tyr	Val	Lys	Ser	Asn	Arg	Leu	Val	Leu	Ala	Thr	Gly	305	310	315	320
Leu	Arg	Asn	Ser	Pro	Gln	Arg	Glu	Ser	Arg	Arg	Lys	Lys	Arg	Gly	Leu	325	330	335	
Phe	Gly	Ala	Ile	Ala	Gly	Phe	Ile	Glu	Gly	Gly	Trp	Gln	Gly	Met	Val	340	345	350	
Asp	Gly	Trp	Tyr	Gly	Tyr	His	His	Ser	Asn	Glu	Gln	Gly	Ser	Gly	Tyr	355	360	365	
Ala	Ala	Asp	Lys	Glu	Ser	Thr	Gln	Lys	Ala	Ile	Asp	Gly	Val	Thr	Asn	370	375	380	
Lys	Val	Asn	Ser	Ile	Ile	Asp	Lys	Met	Asn	Thr	Gln	Phe	Glu	Ala	Val	385	390	395	400
Gly	Arg	Glu	Phe	Asn	Asn	Leu	Glu	Arg	Arg	Ile	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	Leu	Gln	Leu	Arg	Asp	Asn	Ala	450	455	460	

Lys Glu Leu Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn
 465 470 475 480
 Glu Cys Met Glu Ser Ile Arg Asn Gly Thr Tyr Asn Tyr Pro Gln Tyr
 485 490 495
 Ser Glu Glu Ala Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu
 500 505 510
 Glu Ser Ile Gly Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala
 515 520 525
 Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met
 530 535 540
 Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
 545 550 555
 <210> 109
 <211> 548
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> Sản phẩm thiết kế 930
 <400> 109
 Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val
 1 5 10 15
 Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn Leu
 20 25 30
 Leu Glu Asp Ser His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
 35 40 45
 Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
 50 55 60
 Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
 65 70 75 80
 Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly Ser Phe Asn
 85 90 95
 Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
 100 105 110
 Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
 115 120 125
 Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
 130 135 140
 Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
 145 150 155 160
 Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
 165 170 175

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

23591

500	505	510
Leu Ala Ile Tyr Ser Thr Val	Ala Ser Ser Leu Val	Leu Leu Val Ser
515	520	525
Leu Gly Ala Ile Ser Phe Trp	Met Cys Ser Asn Gly	Ser Leu Gln Cys
530	535	540
Arg Ile Cys Ile		
545		
<210>	110	
<400>	110	
000		
<210>	111	
<211>	552	
<212>	PRT	
<213>	Trình tự nhân tạo	
<220>		
<223>	Sản phẩm thiết kế 690 và 734	
<400>	111	
Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val		
1	5	10
Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile		
20	25	30
Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys		
35	40	45
Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn		
50	55	60
Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val		
65	70	75
Glu Lys Ala Asn Pro Thr Asn Asp Leu Cys Tyr Pro Gly His Phe Ala		
85	90	95
Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu		
100	105	110
Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val		
115	120	125
Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr		
130	135	140
Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu		
145	150	155
Ser Lys Ser Tyr Ala Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp		
165	170	175
Gly Val His His Pro Pro Asn Ile Gly Asp Gln Lys Ala Leu Tyr His		
180	185	190

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

23591

Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
 515 520 525
 Leu Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
 530 535 540
 Ser Leu Gln Cys Arg Ile Cys Ile
 545 550
 <210> 112
 <211> 548
 <212> PRT
 <213> Trình tự nhân tạo
 <220>
 <223> Sản phẩm thiết kế 696
 <400> 112
 Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr Val
 1 5 10 15
 Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn Leu
 20 25 30
 Leu Glu Asp Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala
 35 40 45
 Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn
 50 55 60
 Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile Val
 65 70 75 80
 Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Ser Phe Asn
 85 90 95
 Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
 100 105 110
 Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
 115 120 125
 Ser Gly Val Ser Ser Ala Cys Pro Tyr Leu Gly Ser Pro Ser Phe Phe
 130 135 140
 Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Thr Tyr Pro Thr Ile
 145 150 155 160
 Lys Lys Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
 165 170 175
 Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
 180 185 190
 Asn Pro Thr Thr Tyr Ile Ser Ile Gly Thr Ser Thr Leu Asn Gln Arg
 195 200 205
 Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
 210 215 220

23591

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

545

<210> 113

<211> 552

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Sản phẩm thiết kế 691

<400> 113

```

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
1          5          10          15

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
          20          25          30

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile Ala
          35          40          45

Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly Asn
50          55          60

Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile Val
65          70          75          80

Glu Lys Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly His Phe Ala
          85          90          95

Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu
          100          105          110

Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Val
          115          120          125

Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Glu Ser Ser Phe Tyr
130          135          140

Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn Leu
145          150          155          160

Ser Lys Ser Tyr Ala Asn Asn Lys Glu Lys Glu Val Leu Val Leu Trp
          165          170          175

Gly Val His His Pro Pro Asn Ile Gly Asp Gln Lys Ala Leu Tyr His
          180          185          190

Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg Lys
          195          200          205

Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu Gly
210          215          220

Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile Ile
225          230          235          240

Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Arg Tyr Ala Phe Ala Leu
          245          250          255

Ser Arg Gly Phe Gly Ser Gly Ile Ile Asn Ser Asn Ala Pro Met Asp

```

23591

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