



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

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

(11)



1-0032205

(51)⁷ C12N 15/00; A61P 31/16; A61P 37/04; (13) B
C07K 14/005; C12N 7/00; C07K 19/00;
C12N 15/62; C12N 15/82; A61K 39/295;
C07K 14/11

(21) 1-2013-02232

(22) 22/12/2011

(86) PCT/CA2011/001427 22/12/2011

(87) WO2012/083445 28/06/2012

(30) 61/426,401 22/12/2010 US; 61/496,371 13/06/2011 US

(45) 25/06/2022 411

(43) 25/09/2013 306A

(73) MEDICAGO INC. (CA)

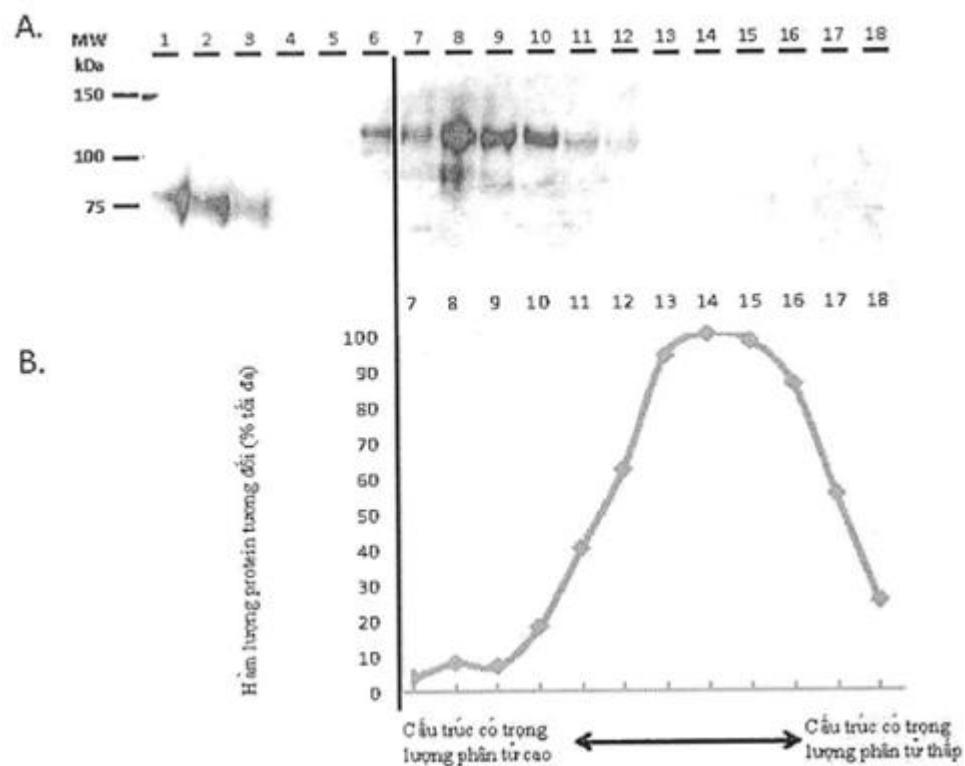
1020, Route de l'Eglise, Suite 600, Québec, Québec G1V 3V9, Canada

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

(74) CÔNG TY LUẬT TRÁCH NHIỆM HỮU HẠN AMBYS HÀ NỘI (AMBYS HANOI)

(54) PHƯƠNG PHÁP SẢN XUẤT HẠT TƯƠNG TỰ VIRUT (VLP) Ở THỰC VẬT VÀ HẠT TƯƠNG TỰ VIRUT ĐƯỢC SẢN XUẤT BẰNG PHƯƠNG PHÁP NÀY

(57) Sáng chế đề cập đến phương pháp sản xuất hạt tương tự virus (virus like particle - VLP) ở thực vật, và chế phẩm bao gồm VLP được tạo ra. Phương pháp bao gồm bước đưa vào axit nucleic bao gồm vùng điều hòa hoạt hóa ở thực vật và được liên kết hiệu quả với trình tự nucleotit thể khảm mã hóa, theo thứ tự, vùng ngoại bào từ protein bề mặt ba đơn phân của virus hoặc đoạn của chúng, được dung hợp vào vùng xuyên màng hoặc đuôi tế bào chất cúm, vào thực vật hoặc bộ phận của thực vật, vùng ngoại bào là từ protein bề mặt ba đơn phân của virus không phải ở bệnh cúm và khác loại về vùng xuyên màng và đuôi tế bào chất cúm. Thực vật hoặc bộ phận của thực vật được ủ dưới các điều kiện cho phép sự biểu hiện axit nucleic, nhờ đó tạo ra VLP. VLP được sản xuất bằng phương pháp này cũng được đề xuất.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc sản xuất protein virus thể khảm ở thực vật. Cụ thể hơn, sáng chế cũng đề cập đến việc sản xuất hạt tương tự virus bao gồm protein virus thể khảm ở thực vật.

Tình trạng kỹ thuật của sáng chế

Việc tiêm chủng tạo ra sự bảo vệ chống lại bệnh gây ra bởi chất tương tự bằng cách làm tăng sự phòng thủ của đối tượng đối với bệnh nhiễm trùng. Thông thường, việc này có thể được tiến hành thông qua việc sử dụng các dạng sống được giảm độc lực hoặc dạng được bất hoạt hoàn toàn của các chất lây nhiễm dưới dạng chất kháng nguyên. Để tránh mối nguy hiểm khi sử dụng virus nguyên vẹn (như virus chết hoặc virus được giảm độc lực) dưới dạng vaccin, protein virus tái tổ hợp, ví dụ các dưới đơn vị, tùy theo dạng vaccin. Cả peptit và vaccin dưới đơn vị đều phụ thuộc vào nhiều giới hạn có hiệu quả. Vaccin dưới đơn vị có thể thể hiện tính gây miễn dịch kém, do sự gấp nếp không đúng, trình diện kháng nguyên kém, hoặc các sai khác về thành phần carbohydrat và lipid. Vấn đề chính là sự khó khăn khi đảm bảo thể cấu tạo của protein được thiết kế bắt chước protein của kháng nguyên trong môi trường tự nhiên của chúng. Tá được thích hợp và, trong trường hợp peptit, protein chất mang, phải được sử dụng để làm tăng đáp ứng miễn dịch. Ngoài ra vaccin này gây ra đáp ứng thể dịch chủ yếu, và do đó có thể không gây ra tính miễn dịch hiệu quả. Vaccin dưới đơn vị thường không hiệu quả đối với bệnh trong đó virus bị bất hoạt hoàn toàn có thể chứng minh tạo ra được sự bảo vệ.

Các hạt tương tự virus (Virus-like particles - VLPs) là đối tượng được lựa chọn có hiệu quả cả trong các chế phẩm gây miễn dịch. Các VLP rất giống với virion trưởng thành, nhưng chúng không chứa vật liệu hệ gen virus. Do đó, VLP không tái bản trong tự nhiên, làm cho chúng an toàn khi sử dụng làm vaccin. Ngoài ra, VLP có thể được thiết kế để biểu hiện glycoprotein virus trên bề mặt của VLP, là thể cấu tạo sinh lý tự nhiên nhất của chúng. Ngoài ra, do VLP giống với virion nguyên vẹn và là cấu trúc hạt đa trị, VLP có thể có hiệu quả trong việc gây ra sự trung hòa các kháng thể với glycoprotein hơn so với kháng nguyên protein vỏ ngoài.

VLP cho hơn ba mươi virus khác nhau được tạo ra trong hệ thống sâu bọ và động vật có vú với mục đích vắc xin (Noad, R. và Roy, P., 2003, *Trends Microbiol* 11: 438-44). Các nghiên cứu khác nhau đã chứng minh rằng protein cúm tái tổ hợp tự lắp ráp thành VLP trong canh trường tế bào sử dụng plasmid biểu hiện ở động vật có vú hoặc vector baculovirus (ví dụ Gomez-Puertas và cộng sự, 1999, *J. Gen. Virol*, 80, 1635-1645; Neumann và cộng sự, 2000, *J. Virol.*, 74, 547-551; Latham và Galarza, 2001, *J. Virol.*, 75, 6154-6165).

Gomez-Puertas và cộng sự (1999, *J. Gen. Virol.*, 80, 1635-1645) chứng minh sự tạo thành hiệu quả của VLP cúm phụ thuộc vào sự biểu hiện các mức protein virus. Neumann và cộng sự (2000, *J. Virol.*, 74, 547-551) thiết lập hệ thống trên cơ sở plasmid biểu hiện ở động vật có vú để tạo ra các hạt tương tự virus cúm lây nhiễm toàn bộ từ cADN được tách dòng. Latham và Galarza (2001, *J. Virol.*, 75, 6154-6165) thông báo sự tạo thành VLP cúm trong tế bào sâu bọ bị nhiễm baculovirus tái tổ hợp biểu hiện đồng thời các gen HA, NA, M1, và M2. Nghiên cứu này chứng minh rằng protein virion cúm tự lắp ráp dựa trên sự biểu hiện đồng thời trong các tế bào sinh vật nhân chuẩn và protein nền M1 cần thiết cho sự sản xuất VLP.

Trong một số hệ thống biểu hiện khác nhau, bao gồm baculovirus, virus bệnh đậu mùa, tế bào ruồi giấm (DS-2), tế bào Vero và tế bào trần nấm men, sự biểu hiện của Pr^{55gag} từ virus gây suy giảm miễn dịch ở người (human immunodeficiency virus - HIV) dẫn đến sự lắp ráp và giải phóng hạt tương tự virus (virus-like particles - VLPs), tương tự với hình thái học của virion HIV chưa trưởng thành (được xem xét bởi Deml và cộng sự, 2005, *Molecular Immunology* 42: 259-277).

Protein vỏ ngoài HIV gp160 có thể được kết hợp thành VLP có nguồn gốc từ Gag. Tuy nhiên, chỉ số lượng hạn chế protein vỏ ngoài được kết hợp không kể sự biểu hiện cao của Pr^{55gag}. Wang và cộng sự chỉ ra rằng sự thay thế vùng xuyên màng và vùng đuôi tế bào chất (TM/CT) của protein vỏ ngoài HIV bằng các vùng protein vỏ ngoài virus khác, bao gồm hemagglutinin cúm dẫn đến làm tăng sự kết hợp của protein vỏ ngoài thành VLP có nguồn gốc từ Pr^{55gag} (*Journal of Virology*, 2007, 81: 10869-10878). Protein vỏ ngoài HIV thể khảm bao gồm HA TM/CT cũng được thể hiện được kết hợp thành VLP có nguồn gốc từ M1 ở bệnh cúm khi được biểu hiện đồng thời trong các tế bào sâu bọ bằng cách sử dụng hệ thống biểu hiện baculovirus (WO2008/005777).

Sự thâm nhập của virus cúm vào tế bào phụ thuộc vào quy trình nhập bào do thụ thể phụ thuộc HA. Chu trình nhiễm virus cúm được bắt đầu bằng cách gắn protein HA bề mặt virion vào thụ thể tế bào chứa axit sialic (glycoprotein và glycolipid). Protein neuraminidase (neuraminidase-NA) do việc xử lý thụ thể axit sialic. Trong các giới hạn axit của hạt cơ quan nội bào được tiếp nhận chứa virion cúm, protein HA trải qua các thay đổi về thể cấu tạo dẫn đến sự dung hợp của virus và màng tế bào, virus không bọc, và sự giải phóng qua trung gian M2 của protein M1 từ ribonucleoprotein (ribonucleoproteins - RNPs) có liên quan đến nucleocapsit, di trú vào nhân tế bào để tổng hợp ARN virus. Latham và Galarza, 200, J. Virol. 75, 6154- 6165) đưa ra sự tạo thành VLP bệnh cúm trong tế bào sâu bọ bị nhiễm baculovirus tái tổ hợp biểu hiện đồng thời các gen HA, NA, M1, và M2. Ngoài ra, Gomez-Puertas và cộng sự, 2000, J. Virol. 74, 11538-11547) chỉ ra rằng, ngoài hemagglutinin (hemagglutinin - HA), protein nền (M1) của virus cúm cần thiết cho nảy chồi của VLP từ các tế bào sâu bọ. Tuy nhiên, Chen và cộng sự (2007, J. Virol. 81, 7111-7123) chỉ ra rằng M1 không cần thiết đối với sự tạo thành VLP.

Cơ chế nảy chồi được mô tả nhiều nhất sử dụng con đường phân loại protein không bào (vacuolar protein sorting - VPS) của vật chủ (xem Chen và Lamb, Virology 372, 2008). Nhiều virus được bao ngoài được chỉ ra để tương tác với protein trong con đường VPS, cần đến tác dụng của phức hợp phân loại chất chứa trong hồng cầu nội thể cần thiết cho các phức hợp protein vận chuyển (ESCRT) (xem Bảng I trong Chen và Lamb 2008). Vùng protein tổng hợp muộn tương tác với protein trong con đường VPS được tìm thấy trên protein lõi và protein nền của virus, và do đó sự nảy chồi phụ thuộc VPS cần thiết cho sự có mặt của protein nền hoặc protein lõi. Sự palmitoyl hóa đuôi tế bào chất của HA cúm cần thiết cho sự nảy chồi nhưng cơ chế của nó không được hiểu rõ và bao gồm các vùng protein bề mặt khác có liên quan. Các yêu cầu tối thiểu để làm nảy chồi chưa được biết đến và sự có mặt của vùng ngoại bào (ectodomain) trong quy trình không được loại trừ. Ngoài ra, đường VPS ở thực vật chưa được hiểu rõ (xem Schellmann S., và Pimpl P., Current Op Plant Biol 12:670-676, 200).

Sự nảy chồi của bệnh cúm đã biết độc lập với đường VPS. Sự nảy chồi của virus cúm bao gồm đường Rab11 (Bruce và cộng sự, J. Virol 84:5848-5859, 2010). Protein Rab là tế bào dạng móc GTPaza được định vị tại bề mặt của túi vận chuyển bên trong tế bào, và chúng có liên quan đến sự tạo thành túi từ cơ thể của người cho, sự vận

chuyển, sự tiếp giáp và dung hợp với cơ thể của người nhận (Vazquez-Martinez và Malagon *Frontiers in Endocrinology* 2:1-9, 2011). Các thành phần của đường Rab11 được nhận biết ở thực vật. Tuy nhiên, các thành phần vận chuyển ở thực vật được tiến triển dẫn đến các đặc tính đặc biệt trong hệ thống nội màng ở thực vật, bao gồm ví dụ không bào chuyên dụng và lớn, sự di chuyển nhanh các chồng Golgi và cơ quan duy nhất của các bộ phận có chất chứa trong hồng cầu nội thể, và số lượng lớn các Rab GTPaza (Rojo E., và Deneke J., *Plant Phys* 147:1493-1503, 2008). Hạt cúm hoặc sự nảy chồi của virus phụ thuộc vào Rab11 (Bruce và cộng sự, *J. Virol* 84:5848-5859, 2010) nhưng protein hoặc vùng protein tương tác với các protein được kết hợp với Rab11 hoặc Rab11 không được nhận biết. Tuy nhiên, vùng rất nhỏ hoặc vùng của HA có thể cần thiết cho quy trình nảy chồi và sự sản xuất VLP chưa được biết đến.

Ở thực vật, HIV Pr^{55gag} tích lũy ở mức vô cùng thấp nếu không được thiết kế để tích tụ trong lục lạp (Meyers và cộng sự, 2008, *BMC Biotechnology* 8:53; Scotti và cộng sự, 2010, *Planta* 229: 1109-1122). Sự tích tụ trong lục lạp, tuy nhiên, có thể không thích hợp với sự kết hợp của protein vỏ ngoài HIV được gấp nếp chính xác do sự trưởng thành và gấp nếp của protein vỏ ngoài cần đến các biến đổi sau dịch mã đặc trưng cho con đường bài tiết. Rybicki và cộng sự (2010, *Plant Biotechnology Journal* 8: 620-637) lưu ý rằng “...đường như chưa có ai biểu hiện thành công toàn bộ protein HIV Env gp160 hoặc thậm chí cả phần chính của protein này, ở thực vật với hiệu suất thích hợp...”.

Virut bệnh dại (rabies virus - RV) là thành viên trong họ Rhabdoviridae. Giống như hầu hết các thành viên trong họ này, RV không được phân đoạn, virus ARN sợi âm mà mã hệ gen của nó dành cho năm protein virus: ARN polymeraza phụ thuộc ARN (L); nucleoprotein (N); protein được phosphoryl hóa (P); protein nền (M) được định vị trên mặt trong của vỏ ngoài của protein virus; và glycoprotein bề mặt ngoài (G). Dietzschold B và cộng sự (1991), *Crit. Rev. Immunol.* 10: 427-439.

Vaccin bệnh dại trên cơ sở nuôi cấy tế bào hạn chế sự phát triển các chủng bị bất hoạt của virus trong canh trường tế bào. Vaccin này bao gồm virus được phát triển trong canh trường tế bào. Các hướng tiếp cận công nghệ sinh học hiện nay nhằm mục đích biểu hiện gen protein màng của virus bệnh dại để phát triển protein tái tổ hợp an toàn có thể được chuyển thành vaccin hoạt hóa. Sự biểu hiện ổn định của glycoprotein virus bệnh dại được chỉ ra trong tế bào trứng chuột đồng Trung Quốc (Burger và cộng

sự, 1991). Thu được protein được glycosyl hóa, đủ độ dài của 67 K được di trú đồng thời bằng G-protein được phân lập từ các tế bào bị nhiễm virus.

WO/1993/001833 đưa ra việc sản xuất hạt tương tự virus (virus like particles-VLPs) trong hệ thống biểu hiện baculovirus chứa hệ gen ARN bao gồm vùng 3' và vùng đệm bao quanh bởi vỏ protein M bệnh dại và protein M1 bệnh dại. VLP cũng bao gồm vỏ ngoài lipid của protein G bệnh dại.

Virus gây bệnh thủy đậu (Varicella Zoster virus - VZV), đã được biết đến dưới dạng virus herpes 3 ở người (HHV-3), là thành viên trong phân họ virus alphaherpes thuộc họ Herpesviridae của virus. VLP biểu hiện glycoprotein hoặc protein vỏ được tạo ra trước đây từ các thành viên trong họ virus herpes khác nhau. Hạt nhẹ (hạt L) bao gồm protein vỏ được bao ngoài, thu được từ các tế bào bị nhiễm virus bệnh viên da do virus herpes simplex typ 1 (herpes simplex virus type 1 - HSV-1), virus herpes ở ngựa typ 1 (equine herpesvirus type 1 - EHV-1), hoặc virus bệnh dại giả (McLauchlan và Rixon (1992) J. Gen. Virol. 73: 269-276; đơn xin cấp Patent Mỹ Số 5384122). Loại VLP khác nhau, được gọi là các hạt được bao ngoài tái bản ADN tiền virus (pre-viral DNA replication enveloped particles - PREPs), có thể được tạo ra từ các tế bào bị nhiễm HSV-1 với sự có mặt của chất ức chế sự tái bản ADN virus. Hạt L tương tự với PREP về mặt cấu trúc, nhưng chứa chế phẩm protein riêng biệt (Dargan và cộng sự (1995) J. Virol. 69: 4924-4932; Patent Mỹ Số 5994116). VLP lai biểu hiện các đoạn protein gE từ VZV được sản xuất bằng kỹ thuật sử dụng protein p1 được mã hóa bằng gen nhảy Ty ở nấm men (Garcia-Valcarcel và cộng sự (1997) Vaccine 15: 709-719; Welsh và cộng sự (1999) J. Med. Virol. 59: 78-83; Patent Mỹ Số 6060064). US 2011/0008838 mô tả VLP thể khảm bao gồm ít nhất một protein VZV, nhưng không bao gồm protein Ty ở nấm men. VLP thể khảm bao gồm protein lõi virus như M1 bệnh cúm hoặc protein M bệnh dịch tả gà và ít nhất một protein virus gây bệnh thủy đậu (varicella zoster virus - VZV).

Sự lan rộng của coronavirus (CoV) được tiến hóa gần đây gây ra mối đe dọa toàn cầu của bệnh dịch lớn hội chứng cấp tính nặng (severe acute respiratory syndrome - SARS) vào năm 2003 (Kuiken, T. và cộng sự, 2003, *Lancet* 362: 263-270). Như với coronavirus khác, SARS-CoV có hình thái học của các hạt bao ngoài với hình chiếu ngoại biên điển hình, được gọi là “vành” hoặc “nhánh,” bao quanh bề mặt của lõi virus (Ksiazek, T. G. và cộng sự, 2003, *N Engl J Med* 348: 1953-1966; Lin, Y. và cộng sự,

2004, *Antivir Ther* 9: 287-289). Ngoài lõi hạt coronavirut là lớp vỏ ngoài lipit chủ yếu chứa ba protein màng, protein M (màng) dồi dào nhất, protein E (vỏ ngoài) nhỏ, và protein S (nhánh). Tam phân đồng nhất của protein S nói chung tạo thành vành được đề cập ở trên, có liên quan đến sự liên kết của virut với thụ thể vật chủ, sự dung hợp màng với sự tiếp nhận virut, sự lan rộng từ tế bào đến tế bào và tính hướng kích thích mô của coronavirut.

Hệ thống biểu hiện baculovirut được sử dụng để tạo ra VLP SARS (Ho, Y. và cộng sự, 2004, *Biochem Biophys Res Commun* 318: 833-838; Mortola, E. và Roy, P., 2004, *FEBS Lett* 576: 174-178). Tuy nhiên, do các khác biệt về bản chất giữa tế bào sâu bọ và tế bào động vật có vú, VLP được lắp ráp trong tế bào sâu bọ (SF9) thể hiện kích thước đường kính bằng 110nm, lớn hơn nhiều so với 78nm virion SARS-CoV xác thực (Lin, Y. và cộng sự, 2004, và Ho, Y. và cộng sự, 2004). Ngoài ra, tính gây miễn dịch của SARS-VLP trên cơ sở tế bào sâu bọ giữ lại không được nghiên cứu. Nhà nghiên cứu khác cũng cố gắng sử dụng hệ thống biểu hiện động vật có vú để tạo ra VLP SARS (Huang, Y. và cộng sự, 2004, *J Virol* 78: 12557-65). Tuy nhiên, sự giải phóng ngoại bào của VLP không có hiệu quả, và hiệu suất của VLP không thỏa mãn. Ví dụ WO/2005/035556 mô tả hệ thống tạo thành các hạt tương tự virut SARS-CoV (SARS-CoV-virus-like particles - SARS-CoV-VLPs) bao gồm một hoặc nhiều vector tái tổ hợp biểu hiện protein SARS-CoV E, protein SARS-CoV M, và protein SARS-CoV S ở tế bào động vật có vú.

Sự tạo thành các VLP, trong hệ thống bất kỳ, các vị trí yêu cầu đáng kể về cấu trúc của các protein - thay đổi sự trải rộng của protein không có nhiều tác dụng đối với sự biểu hiện của chính polypeptit, tuy nhiên các nghiên cứu về cấu trúc không chứng minh được hiệu quả của các biến đổi này đối với sự tạo thành của các VLP. Sự kết hợp của các vùng và cấu trúc khác nhau của protein được tiến hóa bằng virut và không thể cải thiện thành các biến đổi tương tự mà không làm giảm sự tạo thành VLP.

Để cải thiện VLP dưới dạng vaccin lựa chọn hệ thống biểu hiện khác bên cạnh tế bào sâu bọ và động vật có vú cần được khảo sát. Do đó cần phải đánh giá khả năng của hệ thống biểu hiện ở thực vật để tạo ra protein VLP thể khảm. Cụ thể là, cần phải nhận biết số lượng rất ít các protein virut sẽ lắp ráp thành VLP và để đánh giá hình thái học và tính gây miễn dịch của các VLP.

Bản chất kỹ thuật của sáng chế

Sáng chế đề cập đến việc sản xuất protein virus thể khảm ở thực vật. Cụ thể hơn, sáng chế cũng đề cập đến việc sản xuất hạt tương tự virus bao gồm protein virus thể khảm ở thực vật.

Sáng chế đề xuất phương pháp sản xuất hạt tương tự virus (virus like particle - VLP) ở thực vật bao gồm,

a) đưa vào axit nucleic bao gồm vùng điều hòa hoạt hóa ở thực vật và liên kết có hiệu quả với trình tự nucleotit thể khảm mã hóa, theo đây, vùng ngoại bào từ protein bề mặt ba đơn nhân virus hoặc đoạn của nó, được dung hợp vào vùng xuyên màng và đuôi tế bào chất cúm, vào thực vật, hoặc bộ phận của thực vật, vùng ngoại bào từ protein bề mặt ba đơn nhân không phải virus cúm và khác loại với vùng xuyên màng và đuôi tế bào chất cúm, và

b) ủ thực vật hoặc một bộ phận của thực vật dưới các điều kiện cho phép biểu hiện axit nucleic, nhờ đó tạo ra VLP.

Phương pháp như được mô tả ở trên có thể còn bao gồm bước (c) thu hoạch thực vật và tinh chế VLP. Hơn nữa, VLP không chứa protein lõi hoặc protein nền virus.

Sáng chế đề xuất phương pháp được mô tả ở trên, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus hoặc đoạn của chúng có thể nhận được từ các virus trong họ Retroviridae, Rhabdoviridae, Herpesviridae, Coronaviridae, Paramyxoviridae, Poxviridae hoặc Filoviridae. Vùng ngoại bào từ protein bề mặt thể khảm virus có thể có nguồn gốc ví dụ từ họ Lentivirus, Lyssavirus, Varicellovirus, Coronavirus hoặc Ebolavirus. Vùng protein màng mở rộng vào không gian ngoại bào từ protein bề mặt ba đơn phân của virus có thể có nguồn gốc từ ví dụ, nhưng không giới hạn với, HIV, virus bệnh dại, VZV, RSV, virus SARS, virus Ebola, bệnh sởi, bệnh quai bị, bệnh thủy đậu, virus cự bào, Ebola/Filovirus, virus Herpes, virus Epstein-Barr hoặc Smallpox. Protein bề mặt ba đơn phân của virus và dạng tự nhiên của nó có thể bao gồm vùng ngoại bào và vùng xuyên màng/đuôi tế bào chất, ví dụ nhưng không giới hạn với protein F (RSV, bệnh sởi, bệnh quai bị, bệnh dịch tả gà), protein S (SARS), protein env (HIV), protein G (bệnh dại), glycoprotein vỏ ngoài bao gồm E, B, C, I, H (VZV, cytomegalovirus, virus herpes, virus Epstein-barr), glycoprotein GP (ebola, marburg), hemagglutinin (virus bệnh đậu mùa, virus bệnh đậu bò).

Sáng chế cũng bao gồm phương pháp được mô tả ở trên, trong đó vùng xuyên màng và đuôi tế bào chất cúm thu được từ H5 (A/Indonesia/05/2005) hoặc H3 (A/Brisbane/10/2007). Vùng xuyên màng và đuôi tế bào chất có thể bao gồm trình tự nucleotit xác định trong SEQ ID NO:41, hoặc SEQ ID NO:42.

Sáng chế cũng đề xuất phương pháp như được mô tả ở trên, trong đó bước ử (bước b), axit nucleic vào được biểu hiện tạm thời. Mặt khác, trong bước ủ (bước b), axit nucleic vào có thể được biểu hiện ổn định ở thực vật.

Sáng chế cũng bao gồm phương pháp như được mô tả ở trên, trong đó, trong bước đưa vào (bước a), một hoặc nhiều hơn một axit nucleic bổ sung, được chọn từ nhóm bao gồm trình tự nucleotit mã hóa một hoặc nhiều hơn một protein chaperon, protein kênh proton, chất ức chế proteaza, hoặc hỗn hợp của chúng, được đưa vào thực vật.

Sáng chế đề xuất VLP được sản xuất bằng phương pháp được mô tả ở trên. Protein bề mặt ba đơn phân của virus thể khảm của VLP có thể bao gồm N-glycan đặc trưng ở thực vật, hoặc N-glycan được biến đổi. VLP cũng có thể bao gồm một hoặc nhiều hơn một lipid có nguồn gốc từ thực vật.

Sáng chế bao gồm chế phẩm chứa một lượng hữu hiệu VLP như được mô tả ở trên gây ra đáp ứng miễn dịch, và chất mang được dùng.

Sáng chế đề cập đến việc sản xuất protein bề mặt ba đơn phân của virus thể khảm từ các virus trong họ Retroviridae, Rhabdoviridae, Herpesviridae, Coronaviridae hoặc Filoviridae và tạo ra hạt tương tự virus bao gồm protein bề mặt ba đơn phân của virus thể khảm này ở thực vật.

Ngoài ra, sáng chế đề cập đến việc tạo ra virus làm suy giảm miễn dịch ở người (human immunodeficiency virus - HIV), virus bệnh dại, virus gây bệnh thủy đậu (Varicella Zoster Virus - VZV), virus gây hội chứng hô hấp cấp tính nặng (Severe acute respiratory syndrome - SARS) hoặc protein bề mặt ba đơn phân thể khảm virus Ebola ở thực vật. Sáng chế đề cập đến việc sản xuất hạt tương tự virus thể khảm HIV, Rabies, VZV, SARS và Ebola ở thực vật.

Sáng chế đề xuất phương pháp tạo ra HIV thể khảm, bệnh dại, VZV, SARS hoặc Ebola VLP ở thực vật bao gồm bước đưa vào axit nucleic mã hóa protein virus HIV thể khảm, bệnh dại, VZV, SARS hoặc Ebola liên kết có hiệu quả với vùng điều hòa hoạt hóa ở thực vật, vào thực vật hoặc bộ phận của thực vật, và ủ thực vật hoặc bộ

phần của thực vật dưới các điều kiện cho phép sự biểu hiện axit nucleic, nhờ đó tạo ra HIV thể khảm, bệnh dại, VZV, SARS hoặc Ebola VLP.

Sáng chế còn đề xuất đến VLP bao gồm protein HIV thể khảm, bệnh dại, VZV, SARS hoặc Ebola. VLP có thể được tạo ra bằng phương pháp được đề xuất theo sáng chế. HIV, bệnh dại, VZV, SARS hoặc Ebola VLP cũng có thể được tạo ra trong thực vật.

VLP thể khảm, hoặc VLP, được tạo ra từ protein có nguồn gốc từ HIV, bệnh dại, VZV, SARS hoặc Ebola, phù hợp với sáng chế không bao gồm protein M1. Protein M1 đã biết liên kết với ARN có thể được xem là tạp chất trong chế phẩm VLP. Sự có mặt của ARN là không mong muốn khi thu được sự điều hòa chấp nhận được với sản phẩm kháng nguyên (VLP), do đó chế phẩm VLP thể khảm thiếu ARN có thể là một ưu điểm.

Mặc dù protein HIV Env tự nhiên tích lũy kém ở thực vật, nhưng protein HIV Env thể khảm, vẫn được dung hợp với vùng xuyên màng (transmembrane - TM) và đuôi tế bào chất (cytoplasmic tail - CT) từ HA bệnh cúm tích lũy ở mức cao, và nảy chồi thành HIV VLP khi không có mặt protein lõi hoặc protein nền, ở thực vật.

Bản chất kỹ thuật của sáng chế không cần 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 dưới đây bằng cách tham khảo các hình vẽ kèm theo, trong đó:

Fig.1 thể hiện trình tự nucleotit và amino khác nhau, và sự biểu hiện catxet đối với HIV phù hợp với các phương án khác nhau của sáng chế. Fig.1A thể hiện trình tự axit nucleic liên ứng (SEQ ID NO:1) của HIV ConS Δ CFI (Peptit tín hiệu tự nhiên được in đậm, vùng xuyên màng tự nhiên và vùng tế bào chất được gạch dưới). Fig.1B thể hiện trình tự nucleotit của oligonucleotit IF-ApaI-SpPDI.c (SEQ ID NO:2). Fig.1C thể hiện trình tự nucleotit của oligonucleotit SpPDI-HIV gp145.r (SEQ ID NO:3). Fig.1D thể hiện trình tự nucleotit của oligonucleotit IF-SpPDI-gp145.c (SEQ ID NO:4). Fig.1E thể hiện trình tự nucleotit của oligonucleotit WtdTm-gp145.r (SEQ ID NO:5). Fig.1F thể hiện trình tự nucleotit (SEQ ID NO:6) của catxet biểu hiện số 995, từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc NOS). Vị trí SbfI đã loại bỏ được in đậm. PDISP-HIV ConS Δ CFI được gạch dưới. Fig.1G thể hiện

trình tự axit amin của PDISP-HIV ConS Δ CFI (SEQ ID NO:7). Fig.1H thể hiện sự thể hiện dưới dạng sơ đồ của sản phẩm thiết kế số 995.

Fig.2 thể hiện trình tự axit nucleotit và axit amin khác nhau, và catxet biểu hiện đối với HIV phù hợp với các phương án khác nhau của sáng chế. Fig.2A thể hiện trình tự nucleotit có oligonucleotit IF-H3dTm+gp145.r (SEQ ID NO:8). Fig.2B thể hiện trình tự nucleotit có oligonucleotit Gp145+H3dTm.c (SEQ ID NO:9). Fig.2C thể hiện trình tự nucleotit có oligonucleotit H3dTm.r (SEQ ID NO:10). Fig.2D thể hiện trình tự nucleotit (SEQ ID NO:11) của catxet biểu hiện số 997, từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc NOS). Vị trí SbfI được loại trừ được in đậm. PDISP-HIV Con S Δ CFI-A/Brisbane/10/2007 H3 TM+CT được gạch dưới. Fig.2E thể hiện trình tự axit amin của PDISP-HIV ConS Δ CFI-A/Brisbane/10/2007 H3 TM+CT (SEQ ID NO:12). Fig.2F thể hiện sự biểu hiện dưới dạng sơ đồ của sản phẩm thiết kế số 997.

Fig.3. thể hiện trình tự axit nucleotit và axit amin khác nhau, và catxet biểu hiện đối với HIV phù hợp với các phương án khác nhau của sáng chế. Fig.3A thể hiện trình tự nucleotit có oligonucleotit IF-H5dTm+gp145.r (SEQ ID NO: 13). Fig.3B thể hiện trình tự nucleotit có oligonucleotit Gp145+H5dTm.c (SEQ ID NO:14). Fig.3C thể hiện trình tự nucleotit có oligonucleotit IF-H5dTm.r (SEQ ID NO:15). Fig.3D thể hiện trình tự nucleotit (SEQ ID NO:16) của catxet biểu hiện số 999, từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc NOS). Vị trí SbfI đã loại bỏ được in đậm. PDISP-HIV ConS Δ CFI-A/Indonesia/5/2005 H5 TM+CT được gạch dưới. Fig.3E thể hiện trình tự axit amin có PDISP-HIV ConS Δ CFI-A/Indonesia/5/2005 H5 TM+CT (SEQ ID NO:17). Fig.3F thể hiện sự thể hiện dưới dạng sơ đồ của sản phẩm thiết kế số 999.

Fig.4 thể hiện trình tự axit amin và catxet biểu hiện khác nhau với p19 phù hợp với các phương án khác nhau của sáng chế. Fig.4A thể hiện trình tự nucleotit (SEQ ID NO:18) của catxet biểu hiện số 172, từ XmaI (phía trước trình tự khởi đầu plastoxyanin) đến EcoRI (ngay sau trình tự kết thúc plastoxyanin). Trình tự axit nucleic TBSV P19 được gạch dưới. Fig.4B thể hiện trình tự axit amin (SEQ ID NO:19) có gen ức chế TBSV P19 bị bất hoạt. Fig.4C thể hiện sự thể hiện sản phẩm thiết kế số 172.

Fig.5 thể hiện sự thể hiện dưới dạng sơ đồ các gen HIV Env thể khả được biểu hiện như được mô tả dưới đây.

Fig.6 thể hiện phân tích thẩm tách Western của sự biểu hiện protein HIV Env trong các lá *Nicotiana benthamiana* được biểu hiện tạm thời qua trung gian *Agrobacterium*. Làn 1 đến 4, HIV-1 gp160 tái tổ hợp (ab68171), 100, 50, 10 và 5ng, tương ứng, trong 20µg protein lá được chiết từ cây được thâm nhiễm giả (đối chứng dương). Làn 5, 20µg protein tạo thành cây được thâm nhiễm giả (đối chứng âm). Làn 6 đến 8, protein được chiết từ lá được thâm nhiễm AGL1/995 (20µg, 10µg và 2µg, tương ứng, được bổ sung 20µg protein lá được chiết từ cây được thâm nhiễm giả). Làn 9 và 10, protein được chiết từ lá được thâm nhiễm AGL1/997 (20µg và 10µg, tương ứng, được bổ sung 20µg protein lá được chiết từ cây được thâm nhiễm giả). Làn 11 và 12, protein được chiết từ lá được thâm nhiễm AGL1/999 (20µg và 10µg, tương ứng, được bổ sung 20µg protein lá được chiết từ cây được thâm nhiễm giả).

Fig.7 thể hiện sự mô tả các cấu trúc có nguồn gốc từ HIV ConS ΔCFI bằng phương pháp sắc ký loại trừ theo kích thước. Protein chiết từ lá được thâm nhiễm bằng AGL1/999, tạo ra protein khả Env/H5, được tách ra bằng phương pháp lọc gel trên cột phân giải cao S-500 được hiệu chỉnh. (A) Hàm lượng protein HIV Env của phần rửa giải được phát hiện bằng phương pháp phát hiện miễn dịch sử dụng các kháng thể kháng gp120. Làn 1 đến 4, HIV-1 gp160 tái tổ hợp (ab68171), 100, 50, 10 và 5ng, tương ứng, trong 20µg protein lá được chiết từ cây được thâm nhiễm giả (đối chứng dương). Làn 5, 20µg protein tạo thành cây được thâm nhiễm giả (đối chứng âm). Làn 6, 20µg protein được chiết từ lá được thâm nhiễm AGL1/999. Làn 7 đến 18, phần rửa giải từ 7 đến 18 từ phương pháp sắc ký lọc gel. (B) Hàm lượng protein tương đối của các phần rửa giải 7 đến 18 từ phương pháp sắc ký lọc gel.

Fig.8 thể hiện trình tự nucleotit và amino khác nhau, và catxet biểu hiện ở bệnh đại phù hợp với các phương án khác nhau theo sáng chế. Fig.8A thể hiện sơ đồ về sản phẩm thiết kế 1074 (C-2X35S-Rabies glycoprotein G (RabG)+vùng xuyên màng H5 A/Indonesia/5/2005 và đuôi tế bào chất (TM+CT)-NOS trên vector chứa chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin). Fig.8B thể hiện môi IF-RabG-S2+4.c (SEQ ID NO:32). Fig.8C thể hiện môi RabG+H5dTm.r (SEQ ID NO:33). Fig.8D thể hiện gen Rab G được tổng hợp (tương ứng với nt 3317-4891 từ số truy cập ngân hàng gen EF206707; peptit tín hiệu tự nhiên được in đậm, vùng xuyên màng tự nhiên và vùng tế

bào chất được gạch dưới; SEQ ID NO:34). Fig.8E thể hiện môi IF-H5dTm.c (SEQ ID NO:35). Fig.8F thể hiện sản phẩm thiết kế 141 từ t-ADN từ trái sang phải (được gạch dưới). Sự biểu hiện 2X35S-CPMV-HT-PDISP-NOS với catxet biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin (SEQ ID NO:36). Fig.8G thể hiện sự thể hiện dưới dạng sơ đồ sản phẩm thiết kế 141. Vị trí enzyme giới hạn SbfI và StuI được sử dụng để tuyến tính hóa plasmit được giải thích dựa trên sơ đồ biểu diễn. Fig.8H thể hiện trình tự nucleotit của catxet biểu hiện số 1074 từ trình tự khởi đầu 2X35S đến trình tự kết thúc NOS.PDISP-Rab G -A/Indonesia/5/2005 H5 TM+CT được gạch dưới (SEQ ID NO:37). Fig.8I thể hiện trình tự axit amin của PDISP-Rab G-A/Indonesia/5/2005 H5 TM+CT (SEQ ID NO:38). Fig.8J thể hiện trình tự nucleotit của sản phẩm thiết kế 144 từ trái sang phải t-ADN (gạch dưới). 2X35S-CPMV-HT-PDISP-NOS vào hệ thống biểu hiện catxet tái bản + BeYDV với Plastoxyanin-P19-Catxet biểu hiện chất ức chế bất hoạt plastoxyanin (SEQ ID NO:39). Fig.8K thể hiện sự thể hiện dưới dạng sơ đồ sản phẩm thiết kế 144. Vị trí enzyme giới hạn SbfI và StuI được sử dụng để tuyến tính hóa plasmit được giải thích theo sự thể hiện. Fig.8L thể hiện trình tự nucleotit của catxet biểu hiện số 1094 từ phải sang trái BeYDV LIR.PDISP-Rab G -A/Indonesia/5/2005 H5 TM+CT được gạch dưới (SEQ ID NO:40). Fig.8M thể hiện sự thể hiện dưới dạng sơ đồ sản phẩm thiết kế số 1094. Fig.8N thể hiện thử nghiệm thẩm tách miễn dịch về sự biểu hiện protein G bệnh dại ở thực vật. Protein G bệnh dại được biểu hiện khi dung hợp với PDI Sp (sản phẩm thiết kế 1071), BeYDV + rep, vùng H5A/ Indo TM+CT hoặc hỗn hợp của nó. Cụ thể hơn, sản phẩm thiết kế 1074 là sự dung hợp protein G bệnh dại với vùng PDI Sp và H5A/ Indo TM+CT. Sản phẩm thiết kế 1094 là sự dung hợp của protein G bệnh dại với vùng BeYDV + rep, PDI Sp và H5A/ Indo TM+CT. Sản phẩm thiết kế 1091 là sự dung hợp của protein G bệnh dại với PDI Sp và BeYDV + rep. Fig.8O thể hiện thử nghiệm thẩm tách miễn dịch của phương pháp sắc ký loại trừ theo kích thước trên các dịch chiết protein được lọc và cô đặc được biểu hiện từ sản phẩm thiết kế 1074 và sản phẩm thiết kế 1094.

Fig.9A thể hiện thử nghiệm thẩm tách miễn dịch của protein G bệnh dại được tinh chế được biểu hiện từ sản phẩm thiết kế 1074. Fig.9B thể hiện ảnh hiển vi truyền điện tử của VLP được tinh chế từ sự biểu hiện sản phẩm thiết kế 1074.

Fig.10 thể hiện trình tự để điều chế vùng xuyên màng và đuôi tế bào chất A-2X35S-Varicella Zoster Virus glycoprotein E (VZVgE)+H5 A/Indonesia/5/2005 (TM+CT)-NOS (Sản phẩm thiết kế số 946). Fig.10A thể hiện môi IF-wtSp-VZVgE.c (SEQ ID NO 20); Fig.10B thể hiện môi IF-H5dTm+VZVgE.r (SEQ ID NO: 21). Fig.10C thể hiện gen VZV gE được tổng hợp (SEQ ID NO:22; tương ứng với nt 3477-5348 từ ngân hàng gen số truy cập AY013752.1) (Peptit truyền tín hiệu tự nhiên được in đậm, vùng xuyên màng tự nhiên và vùng tế bào chất được gạch dưới). Fig.10D thể hiện môi với VZVgE+H5dTm.c (SEQ ID NO:23). Fig.10E thể hiện catxet biểu hiện số 946 (SEQ ID NO:24), từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc NOS).VZV gE-A/Indonesia/5/2005 H5 TM+CT được gạch dưới. Fig.10F thể hiện trình tự axit amin của VZV gE- A/Indonesia/5/2005 H5 TM+CT (SEQ ID NO:25). Fig.10G thể hiện sự thể hiện dưới dạng sơ đồ của sản phẩm thiết kế số 946

Fig.11A thể hiện thử nghiệm Thẩm tách miễn dịch biểu hiện protein E của virus gây bệnh thủy đậu (Varicella Zoster Virus - VZV). Làn 1 đến 5, VZV gE tái tổ hợp, 500, 100, 50, 10 và 5ng, tương ứng (đối chứng dương). Làn 6 dịch chiết từ lá được thâm nhiễm giả (đối chứng âm). Làn 7-9 protein tái tổ hợp từ sản phẩm thiết kế 946, 20, 10 và 2µg dịch chiết tương ứng. Sản phẩm thiết kế 946 bao gồm gen VZV gE với peptit tín hiệu loại đại và vùng H5A/ Indo TM+CT. Fig.11B thể hiện thử nghiệm thẩm tách miễn dịch để biểu hiện phương pháp sắc ký loại trừ theo kích thước trên dịch chiết thô (sản phẩm thiết kế 946).

Fig.12 thể hiện trình tự nucleotit và amino khác nhau, và catxet biểu hiện với SARS phù hợp với các phương án khác nhau của sáng chế. Fig.12A thể hiện sơ đồ của sản phẩm thiết kế số 916 (B-2X35S- glycoprotein S virus hội chứng hô hấp cấp nghiêm trọng (Severe Acute Respiratory Syndrome Virus glycoprotein S - SARS gS)+vùng xuyên màng và đuôi tế bào chất H5 A/Indonesia/5/2005 (TM+CT)-NOS). Fig.12B thể hiện môi IF-wtSp-SARSGS.c (SEQ ID NO:26). Fig.12C thể hiện môi IF-H5dTm+SARSGS.r (SEQ ID NO:27). Fig.12D thể hiện gen SARS gS được tổng hợp (SEQ ID NO:28; tương ứng với nt 21492-25259 từ số truy cập ngân hàng gen AY278741.1; peptit tín hiệu tự nhiên được in đậm, vùng xuyên màng tự nhiên và vùng tế bào chất được gạch dưới). Fig.12E thể hiện môi SARSGS+H5dTm.c SEQ ID NO:29). Fig.12F thể hiện trình tự nucleotit của catxet biểu hiện số 916 (SEQ ID NO:30), từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc

NOS).SARS gS-A/Indonesia/5/2005 H5 TM+CT được gạch chân. Fig.12G thể hiện trình tự axit amin của SARS gS- A/Indonesia/5/2005 H5 TM+CT (SEQ ID NO:31). Fig.12H thể hiện thử nghiệm thẩm tách miễn dịch của sự biểu hiện của protein S virus hội chứng hô hấp cấp nghiêm trọng (Severe acute respiratory syndrome - SARS). Sản phẩm thiết kế 916 bao gồm gen SARS gS với peptit tín hiệu loại dài và vùng xuyên màng H5 A/Indo và đuôi tế bào chất. Làn 1 dịch chiết từ cây được thâm nhiễm giả (đối chứng âm). Làn 2 đến 4, protein tái tổ hợp từ sản phẩm thiết kế 916 (20, 10 và 5µg dịch chiết).

Fig.13 thể hiện trình tự nucleotit và amino khác nhau, và catxet biểu hiện đối với Ebola phù hợp với các phương án khác nhau của sáng chế. Fig.13A thể hiện trình tự nucleotit của môi IF-Opt_EboGP.s2+4c (SEQ ID NO:43). Fig.13B thể hiện trình tự nucleotit của môi H5iTMCT+Opt_EboGP.r (SEQ ID NO:44). Fig.13C thể hiện trình tự nucleotit của gen GP được tổng hợp tối ưu hóa (tương ứng với nt 6039-8069 từ số truy cập ngân hàng gen AY354458 với trình tự gen loại dài. Trình tự (SEQ ID NO:45) được tối ưu hóa để sử dụng codon và hàm lượng GC, sự xóa bỏ các vị trí nổi bí ẩn, trình tự Shine-Delgarno, trình tự làm mất ổn định ARN và vị trí vào của ribosom của sinh vật nhân nguyên thủy; peptit tín hiệu tự nhiên được in đậm, vùng xuyên màng và vùng tế bào chất được gạch dưới). Fig.13D thể hiện trình tự nucleotit của môi opt_EboGP+H5iTMCT.c (SEQ ID NO: 46). Fig.13E thể hiện sự thể hiện dưới dạng sơ đồ của sản phẩm thiết kế các vị trí enzym giới hạn 1192.SacII và StuI được sử dụng để tuyến tính hóa plasmit được giải thích dựa trên sơ đồ biểu diễn. Fig.13F thể hiện trình tự nucleotit của sản phẩm thiết kế 1192 đường biên t-ADN từ trái sang phải (được gạch dưới).2X35S/CPMV-HT/PDISP/NOS với catxet biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin (SEQ ID NO. 47). Fig.13G thể hiện trình tự nucleotit với catxet biểu hiện số 1366 từ trình tự khởi đầu 2X35S đến trình tự kết thúc NOS.PDISP-GP từ trình tự Zaire 95 Ebolavirus-H5 TM+CT from A/Indonesia/5/2005 được gạch dưới (SEQ ID NO:48). Fig.13H thể hiện trình tự axit amin có PDISP-GP từ Zaire 95 Ebolavirus-H5 TM+CT từ A/Indonesia/5/2005 9SEQ ID NO. 49). Fig.13I thể hiện sự thể hiện dưới dạng sơ đồ sản phẩm thiết kế số 1366. Fig.13J thể hiện phép thử nghiệm thẩm tách miễn dịch của sự biểu hiện protein glycoprotein (GP) virus Ebola. Năm trăm nanogram được pha trong 20µg protein được chiết từ cây được thâm nhiễm giả và được nạp dưới dạng đối chứng dương (C+). Hai mươi microgam protein được

chiết từ cây được thâm nhiễm giả được nạp dưới dạng đối chứng âm (C-). Sản phẩm thiết kế 1366 bao gồm gen GP của virus Ebola với peptit tín hiệu loại đại và vùng H5A/ Indo TM+CT. Các số trong dấu ngoặc đơn chỉ ra lượng canh trường vi khuẩn ban đầu được sử dụng trong chế phẩm chứa chất để tiêm chủng để thâm nhiễm. (200) chỉ ra 200ml canh trường được trộn với 2,3L chất đệm thâm nhiễm và (400) chỉ ra rằng 400ml canh trường được trộn với 2,1L chất đệm thâm nhiễm.

Mô tả chi tiết sáng chế

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

Sáng chế đề cập đến các hạt tương tự virus (virus-like particles - VLPs). Cụ thể hơn, sáng chế đề cập đến VLP bao gồm protein virus thể khảm, và phương pháp sản xuất VLP thể khảm ở thực vật. VLP bao gồm protein dung hợp (thể khảm) bao gồm, theo thứ tự, vùng ngoại bào từ protein bề mặt ba đơn phân của virus (protein bề mặt ba đơn phân của virus) hoặc đoạn của nó, được dung hợp vào vùng xuyên màng và vùng đuôi tế bào chất (TM/CT). Vùng ngoại bào từ protein bề mặt thể khảm virus khác loại với TM/CT. TM/CT là TM/CT từ hemagglutinin cúm (hemagglutinin - HA).

Protein bề mặt ba đơn phân của virus (cũng được đề cập đến dưới dạng protein bề mặt ba đơn phân của virus) là protein được tìm thấy trên bề mặt của virus bao ngoài có dạng trime (thường là tam phân đồng nhất) và bao gồm vùng xuyên màng và vùng đuôi tế bào chất (TM/CT) được định vị tại đầu tận cùng C của mỗi monome và vùng ngoại bào được định vị tại vùng đầu tận cùng N của mỗi monome. Protein bề mặt ba đơn phân của virus có thể là glycoprotein hoặc protein vỏ ngoài. Trime là phức hợp đại phân tử được tạo thành bởi ba, thường là protein được liên kết không phải liên kết cộng hóa trị. Không bị ràng buộc bởi lý thuyết, vùng trime hóa của protein có thể quan trọng đối với sự tạo thành các trime này. Do đó monome của protein bề mặt ba đơn phân của virus hoặc đoạn của nó có thể bao gồm vùng trime hóa. Protein bề mặt ba đơn phân của virus hoặc đoạn của nó còn bao gồm vùng ngoại bào. Vùng ngoại bào của protein bề mặt ba đơn phân được phơi nhiễm với môi trường bên ngoài và không bao gồm vùng xuyên màng và vùng đuôi tế bào chất. Như được mô tả ở đây, vùng ngoại bào từ protein bề mặt ba đơn phân của virus không có nguồn gốc từ virus cúm. Nếu đoạn protein bề mặt ba đơn phân của virus được sử dụng như được mô tả ở đây, ưu tiên protein bề mặt ba đơn phân của virus giữ lại khả năng tạo thành trime.

Vùng xuyên màng và đuôi tế bào chất (TM/CT) của protein bề mặt thể khảm virus, TM/CT của protein cúm, hoặc cả hai, có thể được nhận biết dễ dàng bằng cách sử dụng các phương pháp đã biết bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, ví dụ, bằng cách xác định mức độ kỵ nước của trình tự axit amin của protein, ví dụ bằng cách sử dụng chương trình dự đoán xuyên màng (ví dụ hệ thống phân tích chuyên dụng; ExPASy.org, được hoạt động bởi Swiss Institute of Bioinformatics; hoặc Dense Alignment Surface Method, Cserzo M., và cộng sự 1997, *Prot. Eng.* vol. 10, no. 6, 673-676; Lolkema J.S. 1998, *FEMS Microbiol Rev.* 22, no 4, 305-322), bằng cách xác định đặc tính thủy liệu pháp của trình tự axit amin protein (ví dụ, đặc tính thủy liệu pháp Kyte-Doolittle), bằng cách xác định cấu trúc protein ba chiều và nhận biết cấu trúc ổn định về nhiệt động học trong màng (ví dụ xoắn ốc alpha đơn, phức hợp ổn định của đường xoắn ốc alpha xuyên màng, khoang beta xuyên màng, xoắn ốc beta, hoặc cấu trúc bất kỳ khác ổn định nhiệt động học trong màng). Một lần nữa được nhận biết, vùng TM/CT của protein bề mặt thể khảm virus có thể được thay thế bằng vùng xuyên màng và đuôi tế bào chất thu được từ virus cúm như được mô tả ở trên.

VLP thể khảm theo các phương án khác nhau của sáng chế bao gồm protein bề mặt ba đơn phân của virus, hoặc đoạn protein bề mặt ba đơn phân của virus, từ vùng xuyên màng và vùng đuôi tế bào chất (TM/CT) được thay thế bằng TM/CT thu được từ HA cúm. Protein bề mặt ba đơn phân của virus khác loại đối với TM/CT. Protein bề mặt ba đơn phân của virus, hoặc đoạn của nó, có thể được chuyển hóa mà không giới hạn các virus thuộc họ Retroviridae, Rhabdoviridae, Herpesviridae, Coronaviridae, Paramyxoviridae, Poxviridae hoặc Filoviridae. Protein bề mặt ba đơn phân của virus có nguồn gốc ví dụ từ giống Lentivirus, Lyssavirus, Varicellovirus, Coronavirus hoặc Ebolavirus. Protein bề mặt ba đơn phân của virus có thể nhận được từ ví dụ, nhưng không giới hạn với, HIV, virus bệnh dại, VZV, RSV, virus SARS, virus Ebola, bệnh sởi, quai bị, bệnh thủy đậu, virus cự bào, Ebola/Filovirus, virus Herpes, virus Epstein-Barr hoặc bệnh đậu mùa. Protein bề mặt ba đơn phân của virus có thể ví dụ là protein bề mặt ba đơn phân và dạng tự nhiên của nó có thể bao gồm vùng xuyên màng/đuôi tế bào chất, ví dụ nhưng không giới hạn với:

Protein F (paramyxoviridae: RSV, bệnh sởi, bệnh quai bị, bệnh dịch tả gà)

Protein S (coronaviridae: SARS);

protein env (retroviridae: HIV);

Protein G (rhabdoviridae: bệnh dại);

glycoprotein vỏ ngoài như E, B, C, I, H (herpesviridae: VZV, virus cự bào, virus herpes, virus Epstein-barr);

Glycoprotein GP (filoviridae: ebola, marburg);

hemagglutinin (poxviridae: virus variola, virus bệnh đậu mùa).

Các ví dụ không giới hạn với protein bề mặt ba đơn phân của virus khác nhau có thể được sử dụng theo sáng chế được mô tả chi tiết hơn dưới đây.

HIV

Sáng chế đề xuất các VLP bao gồm protein suy giảm miễn dịch ở người (human immunodeficiency - HIV) thể khả, và các phương pháp tạo ra HIV VLP thể khả, protein HIV thể khả bao gồm protein dung hợp ví dụ protein HIV Env và một phần của hemagglutinin cúm (hemagglutinin - HA), như vùng xuyên màng và vùng đuôi tế bào chất (TM/CT).

Sáng chế đề xuất axit nucleic bao gồm trình tự nucleotit mã hóa protein HIV thể khả liên kết hiệu quả với vùng điều hòa hoạt hóa ở thực vật.

Ngoài ra, sáng chế đề xuất phương pháp sản xuất HIV VLP thể khả ở thực vật. Phương pháp bao gồm bước đưa axit nucleic mã hóa protein HIV thể khả liên kết có hiệu quả với vùng điều hòa hoạt hóa ở thực vật, vào thực vật hoặc bộ phận của thực vật, và ủ thực vật hoặc một bộ phận của thực vật dưới các điều kiện cho phép biểu hiện axit nucleic, nhờ đó tạo ra HIV VLP thể khả.

Sáng chế còn đề xuất VLP bao gồm protein HIV thể khả. VLP có thể được tạo ra bằng phương pháp được đề xuất theo sáng chế.

Virut bệnh dại

Sáng chế đề xuất VLP bao gồm protein virus bệnh dại thể khả, và các phương pháp tạo ra VLP bệnh dại thể khả, protein virus bệnh dại thể khả bao gồm protein dung hợp ví dụ với protein G bệnh dại và một phần của hemagglutinin (hemagglutinin-HA) bệnh cúm, như vùng xuyên màng và vùng đuôi tế bào chất (TM/CT).

Sáng chế đề xuất axit nucleic bao gồm trình tự nucleotit mã hóa protein virus bệnh dại thể khả được liên kết hiệu quả với vùng điều hòa hoạt hóa ở thực vật.

Ngoài ra, sáng chế còn đề xuất phương pháp tạo ra VLP virus bệnh dại thể khả ở thực vật. Phương pháp bao gồm bước đưa vào axit nucleic mã hóa protein virus bệnh

dại thể khảm liên kết hiệu quả với vùng điều hòa hoạt hóa ở thực vật, vào thực vật, hoặc bộ phận của thực vật, và ủ thực vật hoặc bộ phận của thực vật dưới các điều kiện cho phép sự biểu hiện axit nucleic, nhờ đó tạo ra VLP virus bệnh đại thể khảm.

Sáng chế còn đề xuất VLP bao gồm protein virus bệnh đại thể khảm. VLP có thể được tạo ra bằng phương pháp được đề xuất theo sáng chế.

VZV

Sáng chế cũng đề cập đến VLP bao gồm protein của virus gây bệnh thủy đậu (Varicella Zoster Virus - VZV) thể khảm, và phương pháp tạo ra VZV VLP thể khảm, protein VZV thể khảm bao gồm protein dung hợp ví dụ với VZV glycoprotein E và một phần của hemagglutinin (hemagglutinin - HA) bệnh cúm, như vùng xuyên màng và vùng đuôi tế bào chất (TM/CT).

Sáng chế đề xuất axit nucleic bao gồm trình tự nucleotit mã hóa protein VZV thể khảm liên kết có hiệu quả với vùng điều hòa hoạt hóa ở thực vật.

Ngoài ra, sáng chế đề xuất phương pháp tạo ra VZV VLP thể khảm ở thực vật. Phương pháp bao gồm bước đưa vào axit nucleic mã hóa protein VZV thể khảm liên kết hiệu quả với vùng điều hòa hoạt hóa ở thực vật, vào thực vật hoặc bộ phận của thực vật, và ủ thực vật hoặc bộ phận của thực vật dưới các điều kiện cho phép biểu hiện axit nucleic, nhờ đó tạo ra VZV VLP thể khảm.

Sáng chế còn đề xuất VLP bao gồm protein VZV thể khảm. VLP có thể được tạo ra bằng phương pháp được đề xuất theo sáng chế.

SARS

Sáng chế cũng đề cập đến VLP bao gồm protein virus của hội chứng hô hấp cấp nghiêm trọng (Severe acute respiratory syndrome - SARS) thể khảm, và phương pháp tạo ra SARS VLP thể khảm, protein SARS thể khảm bao gồm protein dung hợp ví dụ với SARS glycoprotein S và một phần của hemagglutinin (hemagglutinin - HA) cúm, như vùng xuyên màng và vùng đuôi tế bào chất (TM/CT).

Sáng chế đề xuất axit nucleic bao gồm trình tự nucleotit mã hóa protein SARS thể khảm liên kết hiệu quả với vùng điều hòa hoạt hóa ở thực vật.

Ngoài ra, sáng chế đề xuất phương pháp tạo ra SARS VLP thể khảm ở thực vật. Phương pháp bao gồm bước đưa vào axit nucleic mã hóa protein virus SARS thể khảm được liên kết có hiệu quả với vùng điều hòa hoạt tính ở thực vật, vào thực vật hoặc bộ

phần của thực vật, và ủ thực vật hoặc bộ phận của thực vật dưới các điều kiện cho phép biểu hiện axit nucleic, nhờ đó tạo ra SARS VLP thể khảm.

Sáng chế còn đề xuất VLP bao gồm protein virus SARS thể khảm. VLP có thể được tạo ra bằng phương pháp được đề xuất theo sáng chế.

Ebola

Sáng chế cũng đề cập đến VLP bao gồm protein virus Ebola thể khảm, và phương pháp tạo ra Ebola VLP thể khảm, protein virus Ebola thể khảm bao gồm protein dung hợp ví dụ glycoprotein vỏ ngoài virus Ebola và một phần của hemagglutinin (hemagglutinin - HA) bệnh cúm, như vùng xuyên màng và vùng đuôi tế bào chất (TM/CT).

Sáng chế đề xuất axit nucleic bao gồm trình tự nucleotit mã hóa protein virus ebola thể khảm liên kết hiệu quả với vùng điều hòa hoạt hóa ở thực vật.

Ngoài ra, sáng chế đề xuất phương pháp tạo ra Ebola VLP thể khảm ở thực vật. Phương pháp bao gồm bước đưa ra axit nucleic mã hóa protein virus Ebola thể khảm liên kết hiệu quả với vùng điều hòa hoạt hóa ở thực vật, vào thực vật hoặc bộ phận của thực vật, và ủ thực vật hoặc bộ phận của thực vật dưới các điều kiện cho phép biểu hiện axit nucleic, nhờ đó tạo ra VLP virus Ebola thể khảm.

Sáng chế còn đề xuất VLP bao gồm protein virus Ebola thể khảm. VLP có thể được tạo ra bằng phương pháp được đề xuất theo sáng chế.

Sự tạo thể khảm và sự tạo thành VLP

Cả virus viêm miệng do mụn nước (bệnh đại giống rhabdovirus) và virus Herpes Simplex (virus gây bệnh thủy đậu giống virus herpes) nảy chồi theo cách phụ thuộc vào VSP4 (Taylor và cộng sự J. Virol 81:13631-13639, 2007; Crump và cộng sự, J. Virol 81:7380-7387, 2007). Do VSP4 tương tác với vùng protein nền muộn, điều này giải thích rằng protein nền cần thiết cho sự nảy chồi và, là kết quả, để tạo ra VLP. Tuy nhiên, như được mô tả ở đây, bệnh đại và VZV VLP có thể được tạo ra mà không cần đến protein nền khi thay thế vùng TM/CT bằng các protein của HA cúm. Mà không bị ràng buộc bởi lý thuyết, điều này đề xuất rằng sự tạo thể khảm có thể loại trừ sự phụ thuộc vào sự biểu hiện protein nền đối với sự tạo thành VLP ở bệnh đại và VZV.

Vùng ngoại bào như được mô tả ở trên được dung hợp vào vùng xuyên màng và vùng đuôi tế bào chất (TM/CT), để protein bề mặt ba đơn phân của virus khác loại với TM/CT. TM/CT có thể là TM/CT từ hemagglutinin cúm (hemagglutinin - HA), ví dụ

TM/CT từ H5 hoặc H3, ví dụ nhưng không giới hạn với A/Indonesia/5/05 kiểu phụ (H5N1; “H5/Indo”; số truy cập ngân hàng gen ABW06108.1), H5 A/Vietnam/1194/2004(A-Vietnam; số truy cập ngân hàng gen ACR48874.1), H5 A/Anhui/1/2005 (A-Anhui; số truy cập ngân hàng gen ABD28180.1); H3 A/Brisbane/10/2007 (“H3/Bri”; số truy cập ngân hàng gen ACI26318.1), H3 A/Wisconsin/67/2005(A-WCN; số truy cập ngân hàng gen ABO37599.1). TM/CT của đường biên trong các trình tự H3 và H5 khác nhau được mô tả trong WO 2010/148511 (được kết hợp ở đây để tham khảo). Ví dụ, không được xem xét là hạn chế trình tự axit amin đối với TM/CT có thể bao gồm:

H5 (A/Indonesia/05/2005) TM/CT (SEQ ID NO:41):

QILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI

H3 (A/Brisbane/10/2007) TM/CT (SEQ ID NO:42):

DWILWISFAISCFLLCVALLGFIMWACQKGNIRCNICI

và trình tự nucleotit bất kỳ mã hóa trình tự axit amin có SEQ ID NO:41 hoặc 42.

Trình tự axit nucleic đặc trưng được đề cập đến theo sáng chế, có thể "gần như tương đồng" hoặc “gần như tương tự” với trình tự, hoặc trình tự hoặc phức tạp hóa của trình tự lai giống với một hoặc nhiều hơn một trình tự nucleotit được xác định ở đây dưới các điều kiện lai giống nghiêm ngặt. Trình tự "gần như tương đồng" hoặc “gần như tương tự” khi ít nhất bằng khoảng 70%, hoặc tốt hơn nữa là 75% nucleotit phù hợp với chiều dài xác định của trình tự nucleotit cung cấp trình tự tương đồng này thể hiện một hoặc nhiều hơn một đặc tính của trình tự, hoặc sản phẩm được mã hóa như được mô tả ở đây. Ví dụ vùng ngoại bào gần như tương đồng từ protein bề mặt ba đơn phân virus, được dung hợp với vùng xuyên màng và vùng đuôi tế bào chất thu được từ H3 hoặc H5, vùng xuyên màng và đuôi tế bào chất gần như tương đồng với TM/CT của H3 hoặc H5 và dung hợp với vùng protein màng mở rộng vào không gian ngoại bào từ protein bề mặt ba đơn phân của virus, hoặc cả TM/CT gần như tương đồng và vùng ngoại bào gần như tương đồng từ protein bề mặt ba đơn phân của virus, tạo thành VLP. Sửa chữa sự gấp nếp của protein thể khảm có thể quan trọng đối với sự ổn định của protein, sự tạo thành multimer, tạo thành VLP và chức năng. Sự gấp nếp protein có thể bị ảnh hưởng bởi một hoặc nhiều yếu tố, gồm, nhưng không giới hạn tới, trình tự protein, độ nhiều tương đối của protein, mức độ tập hợp nội bào, tính sẵn có của các

đồng yếu tố có thể liên kết hoặc kết hợp tạm thời với protein không được gấp nếp hoặc được gấp nếp một phần, được gấp nếp.

Tính tương tự của trình tự này có thể được xác định bằng cách sử dụng chương trình so sánh trình tự nucleotit, như được đề xuất trong DNASIS (sử dụng, ví dụ nhưng không giới hạn tới, các thông số dưới đây: GAP penalty 5, #of top diagonals 5, fixed GAP penalty 10, k-tuple 2, floating gap 10, và window size 5). Tuy nhiên, các phương pháp sắp hàng trình tự khác để so sánh đã được biết trong lĩnh vực kỹ thuật ví dụ thuật toán của Smith & Waterman (1981, Adv. Appl. Math. 2:482), Needleman & Wunsch (J. Mol. Biol. 48:443, 1970), Pearson & Lipman (1988, Proc. Nat'l. Acad. Sci. USA 85:2444), và bằng các phép tiến hành bằng máy tính các thuật toán này (GAP, BESTFIT, FASTA, và BLAST, có sẵn thông qua NIH.), hoặc bằng cách sắp hàng thủ công và kiểm tra bằng mắt (xem, ví dụ, Current Protocols in Molecular Biology, Ausubel và cộng sự, eds. 1995 bản phụ lục), hoặc bằng cách sử dụng phép lai Southern hoặc Northern dưới các điều kiện khắc nghiệt (xem Maniatis và cộng sự, trong Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory, 1982). Tốt hơn là, các trình tự gần như tương đồng thể hiện ít nhất khoảng 80% và tốt hơn nhất ít nhất khoảng 90% tính tương tự về trình tự đối với chiều dài xác định của phân tử.

Ví dụ về một trong các điều kiện lai khắc nghiệt có thể lai qua đêm (từ khoảng 16 đến 20 giờ) 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 thời gian một giờ, hoặc rửa 2 lần trong 0,1 X SSC ở nhiệt độ 65°C mỗi lần rửa khoảng 20 hoặc 30 phút. Một khác điều kiện lai khắc nghiệt dẫn chúng có thể là qua đêm (16-20 giờ) trong 50% formamit, 4 X SSC ở nhiệt độ 42°C, sau đó bằng cách rửa trong 0,1 X SSC ở nhiệt độ 65°C trong thời gian một giờ, hoặc rửa 2 lần trong 0,1 X SSC ở nhiệt độ 65°C mỗi lần trong thời gian 20 hoặc 30 phút, hoặc qua đêm (từ 16 đến 20 giờ), hoặc phương pháp lai trong dung dịch đệm nước phosphat Church (7% SDS; 0,5M chất đệm NaPO_4 pH= 7,2; 10mM EDTA) ở nhiệt độ 65°C, với 2 lần rửa ở nhiệt độ 50°C trong 0,1 X SSC, 0,1% SDS mỗi lần trong thời gian 20 hoặc 30 phút, hoặc rửa 2 lần ở nhiệt độ 65°C trong 2 X SSC, 0,1% SDS trong thời gian 20 hoặc 30 phút mỗi lần với các vùng trình tự duy nhất.

Axit nucleic mã hóa polypeptit thể khảm, protein thể khảm, protein dung hợp, protein virus thể khảm, hoặc protein bề mặt ba đơn phân của virus thể khảm có thể

được mô tả dưới dạng "axit nucleic thể khảm", hoặc "trình tự nucleotit thể khảm". Ví dụ, nhưng không được xem là để giới hạn, hạt tương tự virus bao gồm protein HIV thể khảm, protein virus bệnh dại thể khảm, protein VZV thể khảm, SARS thể khảm hoặc protein virus Ebola thể khảm có thể được mô tả dưới dạng "VLP thể khảm".

"Protein khảm" hoặc "polypeptit khảm", cũng được đề cập đến dưới dạng protein dung hợp, nó được hiểu là protein hoặc polypeptit bao gồm trình tự axit amin từ hai hoặc nhiều hơn hai nguồn, ví dụ nhưng không giới hạn với vùng ngoại bào từ protein bề mặt ba đơn phân của virus hoặc đoạn của nó, ví dụ protein F (ví dụ RSV, bệnh sởi, bệnh quai bị, bệnh dịch tả gà), protein (ví dụ SARS), protein env (HIV), protein G (bệnh dại), glycoprotein vỏ ngoài như E, B, C, I, H (VZV, virus cự bào, virus herpes, virus Epstein-barr), glycoprotein GP (ví dụ ebola, marburg), hemagglutinin (ví dụ virus bệnh đậu mùa, virus bệnh đậu bò), và ví dụ TM/CT của HA, được dung hợp thành polypeptit đơn. Protein hoặc polypeptit khảm có thể bao gồm peptit tín hiệu tương tự, hoặc khác loại với, phần còn lại của polypeptit hoặc protein.

Thuật ngữ "peptit tín hiệu" đã được biết trong lĩnh vực kỹ thuật và thường đề cập đến trình tự ngắn (khoảng 5 đến 30 axit amin) của axit amin, thường được tìm thấy tại đầu tận cùng N của polypeptit có thể chuyển đoạn trực tiếp polypeptit được dịch mã mới thành cơ quan tế bào cụ thể, hoặc hỗ trợ sự định vị các vùng đặc hiệu của chuỗi polypeptit tương ứng với các vùng khác. Như ví dụ không giới hạn, peptit tín hiệu có thể hướng đích sự chuyển đoạn của protein vào màng lưới nội chất và/hoặc sự hỗ trợ sự định vị của vùng ở đầu gần đầu tận cùng N tương ứng với vùng màng dạng móc của polypeptit mới sinh để hỗ trợ sự phân cắt và gấp nếp protein mới sinh, ví dụ được xem là được giới hạn, protein HA trưởng thành.

Peptit tín hiệu (signal peptide-SP) có thể tự nhiên với protein hoặc protein virus, hoặc peptit tín hiệu có thể khác loại với trình tự cơ bản của protein hoặc protein virus được biểu hiện. Protein hoặc protein virus có thể bao gồm peptit tín hiệu từ loại cúm thứ nhất, kiểu phụ hoặc chủng cân bằng của HA từ một hoặc nhiều hơn một loại bệnh cúm khác nhau, kiểu phụ hoặc chủng. Ví dụ peptit tín hiệu tự nhiên của HA kiểu phụ B H1, H2, H3, H5, H6, H7, H9 hoặc loại bệnh cúm B có thể được sử dụng để biểu hiện protein virus thể khảm trong hệ thống thực vật. Theo một vài phương án của sáng chế, SP có thể thuộc loại bệnh cúm B, H1, H3 or H5; hoặc thuộc kiểu phụ H1/Bri, H1/NC, H5/Indo, H3/Bri hoặc B/Flo. Theo một vài phương án SP có thể là protein

HIV Env, protein G bệnh dại, glycoprotein E VZV, glycoprotein S SARS hoặc glycoprotein vỏ ngoài virus Ebola.

Peptit tín hiệu cũng có thể không tự nhiên, ví dụ, từ protein, protein virus, protein bề mặt ba đơn phân của virus hoặc hemagglutinin của virus khác với protein bề mặt ba đơn phân của virus, hoặc từ polypeptit thực vật, động vật hoặc vi khuẩn. Ví dụ không giới hạn của peptit tín hiệu có thể được sử dụng là peptit của protein disulfua isomerase ở cỏ linh lăng (PDI SP; nucleotit 32-103 có số truy cập Z11499).

Do đó sáng chế đề xuất protein virus thể khảm bao gồm peptit tín hiệu tự nhiên, hoặc không tự nhiên, và axit nucleic mã hóa protein virus thể khảm này.

Sự gấp nếp chính xác protein virus thể khảm được biểu hiện có thể quan trọng đối với độ ổn định của protein, sự tạo thành multimer, sự tạo thành các VLP, chức năng của protein virus thể khảm và sự nhận diện protein virus thể khảm bởi kháng thể, trong số các đặc tính khác. Sự gấp nếp và tích tụ protein có thể bị ảnh hưởng bởi một hoặc nhiều yếu tố, gồm, nhưng không giới hạn với, trình tự của protein, sự dư thừa protein tương đối, mức độ dồn nội bào, pH trong ngăn tế bào, tính sẵn có của các đồng yếu tố có thể liên kết hoặc kết hợp tạm thời với protein được gấp nếp, được gấp nếp một phần hoặc không được gấp nếp, sự có mặt của một hoặc nhiều protein chaperon, hoặc loại khác.

Protein sốc nhiệt (Heat shock protein-Hsp) hoặc protein gây căng thẳng là các ví dụ về protein chaperon, có thể tham gia vào các quá trình tế bào khác nhau bao gồm sự tổng hợp protein, sự lưu thông nội bào, ngăn chặn sự gấp nếp sai, ngăn chặn sự kết tụ protein, sự lắp ráp và tháo rời các phức hợp protein, sự gấp nếp protein, và sự không kết tụ protein. Các ví dụ về protein chaperon này gồm, nhưng không giới hạn với, Hsp60, Hsp65, Hsp 70, Hsp90, Hsp100, Hsp20-30, Hsp10, Hsp100-200, Hsp100, Hsp90, Lon, TF55, FKBP, cyclophilin, ClpP, GrpE, ubiquitin, calnexin, và protein disulfua isomerase (xem, ví dụ, Macario, A.J.L., *Cold Spring Harbor Laboratory Res.* 25:59-70. 1995; Parsell, D.A. & Lindquist, S. *Ann. Rev. Genet.* 27:437-496 (1993); Patent Mỹ số 5232833). Như được mô tả ở đây, protein chaperon, ví dụ nhưng không giới hạn với Hsp40 và Hsp70 có thể được sử dụng để đảm bảo sự gấp nếp protein virus thể khảm.

Các ví dụ về Hsp70 gồm Hsp72 và Hsc73 từ các tế bào động vật có vú, DnaK từ vi khuẩn, đặc biệt là vi khuẩn gây bệnh lao mycobacteria như *Mycobacterium leprae*,

Mycobacterium tuberculosis, và *Mycobacterium bovis* (như Bacille-Calmette Guerin: được đề cập ở đây có dạng Hsp71). DnaK từ *Escherichia coli*, nấm men và sinh vật nhân nguyên thủy khác, và BiP và Grp78 từ sinh vật nhân chuẩn, như *A. thaliana* (Lin và cộng sự 2001 (Cell Stress và Chaperones 6:201-208). Ví dụ cụ thể về Hsp70 là *A. thaliana* Hsp70 (được mã hóa bằng Genbank ref: AY120747.1). Hsp70 có khả năng liên kết đặc hiệu với ATP cũng như polypeptit và peptit không được gấp nếp, nhờ đó tham gia vào sự gấp nếp và không gấp nếp protein cũng như sự lắp ráp và tháo dỡ các phức hợp protein.

Các ví dụ về Hsp40 gồm DnaJ từ sinh vật nhân nguyên thủy như *E. coli* và mycobacteria và HsJ1, HDJ1 và Hsp40 từ sinh vật nhân chuẩn, như cỏ linh lăng (Frugis và cộng sự, 1999. Plant Molecular Biology 40:397-408). Ví dụ cụ thể về Hsp40 là *M. sativa* MsJ1 (số tham chiếu ngân hàng gen: AJ000995.1). Hsp40 đóng vai trò làm chaperon phân tử đối với sự gấp nếp protein, tính dung nạp nhiệt và sự tái bản ADN, trong số các hoạt tính tế bào khác.

Trong số Hsps, Hsp70 và đồng chaperon của nó, Hsp40, có liên quan đến sự ổn định sự dịch mã và polypeptit được tổng hợp mới trước khi sự tổng hợp hoàn thành. Mà không bị ràng buộc bởi lý thuyết, Hsp40 liên kết với các miếng dán kỵ nước của polypeptit không gấp nếp (mới sinh hoặc được chuyển mới), do đó tạo thuận lợi cho sự tương tác của phức hợp Hsp70-ATP với polypeptit. Sự thủy phân ATP dẫn đến sự tạo thành phức hợp ổn định giữa polypeptit, Hsp70 và ADP, và giải phóng Hsp40. Sự kết hợp của phức hợp Hsp70-ADP với miếng dán kỵ nước chứa polypeptit ngăn chặn sự tương tác của chúng với các miếng dán kỵ nước khác, ngăn chặn sự gấp nếp không chính xác và sự tạo thành các khối kết tụ với protein khác (tham khảo trong Hartl, FU. 1996. Nature 381:571-579).

Protein có thể tạo thuận lợi cho sự gấp nếp chính xác ở mức thấp của protein tái tổ hợp, nhưng dưới dạng mức biểu hiện tăng, sự dư thừa chaperon tự nhiên có thể trở thành một yếu tố hạn chế. Mức biểu hiện cao của protein virus thể khảm trong các lá được thâm nhiễm có thể dẫn đến sự tích tụ của protein virus thể khảm trong dịch bào tương, và biểu hiện đồng thời một hoặc nhiều hơn một protein chaperon như Hsp70, Hsp40 hoặc cả Hsp70 và Hsp40 có thể làm giảm mức protein được gấp nếp sai hoặc được kết tụ, và làm tăng số lượng protein thể hiện các đặc tính cấu trúc bậc bốn và bậc ba cho phép tạo thành các hạt tương tự virus.

Do đó, sáng chế cũng đề xuất phương pháp sản xuất protein virus thể khảm VLP ở thực vật, trong đó axit nucleic thứ nhất mã hóa protein virus thể khảm được biểu hiện đồng thời với axit nucleic thứ hai mã hóa chaperon. Axit nucleic thứ nhất và thứ hai có thể được đưa vào thực vật trong bước tương tự, hoặc có thể được đưa vào thực vật một cách tuần tự.

VLP thể khảm được tạo ra từ protein có nguồn gốc từ virus, phù hợp với sáng chế không bao gồm protein nền hoặc protein lõi của virus. Protein nền virus là protein thiết lập và duy trì cấu trúc virion. Protein nền virus thường tương tác trực tiếp với màng tế bào và có thể bao gồm trong quy trình nảy chồi. Protein lõi virus là protein tạo thành một phần của nucleocapsit và thường được kết hợp trực tiếp với axit nucleic của virus. Các ví dụ về protein nền hoặc protein lõi của virus là M1, RSV M bệnh cúm và protein gag retrovirus. Protein M1 đã biết liên kết với ARN (Wakefield L., và Brownlee G.G., Nucl Acids res 11:8569-8580, 1989) là tạp chất trong chế phẩm VLP. Sự có mặt của ARN không mong muốn khi thu được sự điều hòa chấp nhận được cho sản phẩm VLP thể khảm, do đó chế phẩm VLP thể khảm thiếu ARN có thể là một ưu điểm.

Việc sử dụng thuật ngữ "vùng điều hòa", "nhân tố điều hòa" hoặc "trình tự khởi đầu" trong đơn được hiểu là để phản ánh một phần của axit nucleic thông thường, nhưng không thường xuyên, phía trước vùng mã hóa protein của gen, có thể bao gồm ADN hoặc ARN, hoặc cả ADN và ARN. Khi vùng điều hòa được hoạt hóa, và trong sự kết hợp hiệu quả, hoặc được liên kết 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 làm trung gian tính đặc trưng của cơ quan, hoặc kiểm soát sự hoạt hóa gen tạm thời hoặc phát triển. "Vùng điều hòa" có thể bao gồm nhân tố khởi động, nhân tố khởi động lõi thể hiện hoạt tính khởi động cơ bản, nhân tố có thể gây ra đáp ứng với tác nhân kích thích bên ngoài, các nhân tố qua trung gian hoạt tính khởi động như nhân tố điều hòa âm tính hoặc vùng tăng cường phiên mã. "Vùng điều hòa", như được sử dụng ở đây, cũng có thể bao gồm nhân tố hoạt hóa sự phiên mã sau đó, ví dụ, nhân tố điều hòa điều biến sự biểu hiện gen như vùng tăng cường dịch mã và phiên mã, chất kìm hãm dịch mã và phiên mã, phía trước trình tự hoạt hóa, và yếu tố quyết định không ổn định của mRNA. Một trong số các nhân tố này có thể được đặt vào vị trí gần vùng mã hóa.

Trong ngữ cảnh của bản mô tả, thuật ngữ "nhân tố điều hòa" hoặc "vùng điều hòa" thường dùng để chỉ trình tự ADN, thường, nhưng không luôn luôn, phía trước (5') trình tự mã hóa của gen cấu trúc, kiểm soát sự biểu hiện của vùng mã hóa bằng cách cung cấp sự nhận diện ARN polymeraza và/hoặc yếu tố khác cần thiết cho sự phiên mã để bắt đầu tại vị trí đặc biệt. 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. Hầu hết, nhưng không phải là tất cả, các nhân tố trình tự khởi đầu 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 phía trước vị trí khởi đầu phiên mã. Nhân tố trình tự khởi đầu bao gồm nhân tố trình tự khởi đầu 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 vài vùng điều hòa được điều chỉnh sự phát triển có thể hoạt hóa ưu tiên trong các cơ quan hoặc mô nhất định tại các giai đoạn phát triển đặc biệt, chúng có thể được hoạt hóa theo cách thức điều hòa sự phát triển, hoặc với mức độ cơ bản trong các cơ quan hoặc mô khác trong thực vật. Các ví dụ về vùng điều hòa đặc hiệu mô, ví dụ vùng điều hòa đặc hiệu, bao gồm trình tự khởi đầu napin, và trình tự khởi đầu cruciferin (Rask và cộng sự, 1998, J. Plant Physiol. 152: 595-599; Bilodeau và cộng sự., 1994, Plant Cell 14: 125-130). Một ví dụ về trình tự khởi đầu đặc hiệu lá bao gồm trình tự khởi đầu plastoxyanin (xem US 7125978, được kết hợp ở đây để tham khảo).

Vùng điều hòa cảm ứng được là một vùng có khả năng hoạt hóa sự phiên mã trực tiếp hoặc gián tiếp một hoặc nhiều trình tự ADN hoặc các gen đáp ứng 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 cảm ứng được 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 không có 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 được gây ra trực tiếp bởi các nhân tố nóng, lạnh, muối, hoặc độc tố hoặc gián tiếp thông qua hoạt động của mầm bệnh hoặc tác nhân có khả năng gây bệnh như virus. Tế bào thực vật chứa vùng điều hòa có thể cảm ứng được phơi ra với tác nhân gây cảm ứng bằng cách đặt bên ngoài tác nhân gây cảm ứng với tế bào hoặc thực vật như bằng cách phun, tưới nước, làm nóng hoặc các phương pháp tương tự. Nhân tố điều hòa có thể cảm ứng có thể có nguồn gốc từ các gen của thực vật hoặc các gen không phải của thực vật (ví dụ Gatz, C. và Lenk, L.R.P., 1998, Trends Plant Sci. 3, 352-358; được kết hợp ở đây để tham khảo). Ví dụ, trình tự khởi đầu có thể cảm ứng hiệu quả gồm, nhưng không giới hạn với, trình tự khởi đầu có thể cảm ứng với tetracyclin (Gatz, C., 1997, Ann. Rev. Plant Physiol. Plant Mol. Biol. 48, 89-108; được kết hợp để tham khảo), trình tự khởi đầu có thể cảm ứng steroid (Aoyama, T. và Chua, N.H., 1997, Plant 1. 2, 397-404; được kết hợp để tham khảo) và trình tự khởi đầu có thể cảm ứng với rượu etylic (Salter, M.G., và cộng sự, 1998, Plant Journal 16, 127-132; Caddick, M.X., và cộng sự, 1998, Nature Biotech. 16, 177-180, được kết hợp để tham khảo) gen IB6 và CKI 1 có thể cảm ứng xytokinin (Brandstatter, I. và Kieber, J., 1998, Plant Cell 10, 1009-1019; Kakimoto, T., 1996, Science 274, 982-985; được kết hợp để tham khảo) và nhân tố có thể cảm ứng với auxin, DR5 (Ulmasov, T., và cộng sự, 1997, Plant Cell 9, 1963-1971; được kết hợp để tham khảo).

Vùng điều hòa cơ định nhằm biểu hiện gen trong toàn bộ các phần khác nhau của thực vật và trong toàn bộ sự phát triển liên tục của thực vật. Các ví dụ về nhân tố điều hòa cấu thành gồm trình tự khởi đầu có liên quan đến bản phiên mã CaMV 35S (Odell và cộng sự, 1985, Nature, 313: 810-812), actin 1 ở gạo (Zhang và cộng sự, 1991, Plant Cell, 3: 1155-1165), actin 2 (An và cộng sự, 1996, Plant J., 10: 107-121), hoặc tms 2 (U.S. 5,428,147, được kết hợp ở đây để tham khảo), và gen triosephosphat isomerase 1 (Xu và cộng sự, 1994, Plant Physiol. 106: 459-467), gen ubiquitin 1 ở ngô (Cornejo và cộng sự, 1993, Plant Mol. Biol. 29: 637-646), gen ubiquitin 1 và 6 ở *Arabidopsis*

(Holtorf và cộng sự, 1995, Plant Mol. Biol. 29: 637-646), và gen 4A yếu tố khởi động sự dịch mã ở thuốc lá (Mandel và cộng sự, 1995, Plant Mol. Biol. 29: 995-1004).

Thuật ngữ "cơ định" được sử dụng trong bản mô tả không cần thiết phải chỉ ra gen nhờ sự kiểm soát vùng điều hòa cơ định được biểu hiện tại mức giống như trong tất cả các loại tế bào, nhưng gen được biểu hiện trong khoảng rộng các loại tế bào mặc dù thường quan sát thấy nhiều thay đổi. Nhân tố điều hòa cơ định có thể được kết hợp với các trình tự khác để làm tăng cường sự phiên mã và/hoặc dịch mã trình tự nucleotit mà chúng được liên kết hiệu quả. Ví dụ, hệ thống CPMV-HT nhận được từ vùng không dịch mã của virus khảm ở cây đậu đũa (Cowpea mosaic virus - CPMV) và chứng minh sự chuyển đoạn tăng cường của trình tự mã hòa được kết hợp. "Tự nhiên" được hiểu là axit nucleic hoặc trình tự axit amin xuất hiện tự nhiên, hoặc "loại đại". Nhờ "được liên kết 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.

Protein hoặc polypeptit thể khảm có thể được biểu hiện trong hệ thống biểu hiện bao gồm virus trên cơ sở, ADN hoặc ARN, hệ thống biểu hiện, ví dụ nhưng không giới hạn với, catxet biểu hiện trên cơ sở comovirus và nhân tố phóng đại trên cơ sở geminivirus.

Hệ thống biểu hiện được mô tả ở đây có thể bao gồm catxet biểu hiện trên cơ sở virus hai phần, hoặc virus với hệ gen hai phần. Ví dụ, virus hai phần có thể thuộc họ Comoviridae. Giống thuộc họ Comoviridae bao gồm Comovirus, Nepovirus, Fabavirus, Cheravirus và Sadwavirus. Comovirus bao gồm virus khảm ở cây đậu đũa (Cowpea mosaic virus - CPMV), virus khảm xấu ở cây đậu đũa (Cowpea severe mosaic virus - CPSMV), virus khảm ở bí (Squash mosaic virus - SqMV), virus đốm ở cỏ ba lá đỏ (Red clover mottle virus - RCMV), virus đốm ở vỏ đậu (Bean pod mottle virus - BPMV), virus đốm ở củ cải (Turnip ringspot virus - TuRSV), virus khảm ở đậu tằm (Broad bean true mosaic virus - BBtMV), virus làm biến màu ở đậu tằm (Broad bean stain virus - BBSV), virus khảm ở củ cải (Radish mosaic virus - RaMV). Các ví dụ về trình tự ARN-2 ở comovirus bao gồm nhân tố tăng cường có thể hữu ích trong nhiều khía cạnh khác nhau của sáng chế, gồm, nhưng không giới hạn với: CPMV RNA-2 (số

truy cập ngân hàng gen NC_003550), RCMV RNA-2 (số truy cập ngân hàng gen NC_003738), BPMV RNA-2 (số truy cập ngân hàng gen NC_003495), CPSMV RNA-2 (số truy cập ngân hàng gen NC_003544), SqMV RNA-2 (số truy cập ngân hàng gen NC_003800), TuRSV RNA-2 (số truy cập ngân hàng gen NC_013219.1). BBtMV RNA-2 (số truy cập ngân hàng gen GU810904), BBSV RNA2 (số truy cập ngân hàng gen FJ028650), RaMV (số truy cập ngân hàng gen NC_003800)

Các đoạn của hệ gen ARN comovirut hai phần được đề cập đến dưới dạng ARN-1 và ARN-2. ARN-1 mã hóa protein có liên quan đến sự tái bản trong khi ARN-2 mã hóa protein cần thiết cho sự dịch chuyển từ tế bào đến tế bào và hai protein vỏ. catxet trên cơ sở comovirut thích hợp có thể được sử dụng bao gồm CPMV, CPSMV, SqMV, RCMV, hoặc BPMV, ví dụ, catxet biểu hiện có thể trên cơ sở CPMV.

“Catxet biểu hiện” dùng để chỉ trình tự nucleotit bao gồm axit nucleic đang nói đến nhờ kiểm soát, và liên kết hiệu quả (hoặc một cách hiệu quả), trình tự khởi đầu thích hợp hoặc nhân tố điều hòa khác để phiên mã của axit nucleic đang nói đến trong tế bào chủ.

Các hệ thống biểu hiện cũng có thể bao gồm nhân tố phóng đại từ geminivirut ví dụ, nhân tố phóng đại từ virut gây lùn ở đậu vàng (bean yellow dwarf virus - BeYDV). BeYDV thuộc giống Mastreviruses phù hợp với cây hai lá mầm. BeYDV là hai phần có hệ gen ADN vòng tròn chuỗi đơn và có thể tái bản thành số lượng bản sao rất lớn bằng cơ chế tuần hoàn lẫn. Hệ thống vector tái bản ADN có nguồn gốc từ BeYDV được sử dụng để sản xuất protein hiệu suất cao nhanh chóng ở thực vật.

Như được sử dụng ở đây, cụm từ “nhân tố phóng đại” dùng để chỉ đoạn axit nucleic bao gồm ít nhất một phần của một hoặc nhiều vùng hai gen dài (long intergenic region - LIR) hoặc hệ gen geminivirut. Như được sử dụng ở đây, “vùng hai gen dài” dùng để chỉ vùng trong vùng hai gen dài chứa vị trí liên kết rep có khả năng cắt bỏ gián tiếp và tái bản bằng protein Rep geminivirut. Theo một vài khía cạnh, đoạn axit nucleic bao gồm một hoặc nhiều LIR, còn bao gồm vùng hai gen ngắn (short intergenic region - SIR) của hệ gen geminivirut. Như được sử dụng ở đây, “vùng hai gen ngắn” dùng để chỉ sợi bổ sung (IR ngắn (SIR) của Mastrevirut). Nhân tố phóng đại có nguồn gốc từ geminivirut thích hợp bất kỳ có thể được sử dụng ở đây. Xem, ví dụ, WO2000/20557; WO2010/025285; Zhang X. và cộng sự (2005, *Biotechnology and Bioengineering*, Vol. 93, 271-279), Huang Z. và cộng sự (2009, *Biotechnology and*

Bioengineering, Vol. 103, 706-714), Huang Z. và cộng sự(2009, *Biotechnology and Bioengineering*, Vol. 106, 9-17); được kết hợp ở đây để tham khảo).

Protein khảm hoặc polypeptit khảm có thể được tạo ra dưới dạng sản phẩm phiên mã từ trình tự nucleotit khảm, và protein khảm hoặc polypeptit khảm được tách ra sau khi tổng hợp, và nếu cần, được kết hợp để tạo thành protein multime. Do đó, protein khảm hoặc polypeptit khảm cũng bao gồm protein hoặc polypeptit bao gồm các dưới đơn vị được kết hợp thông qua các cầu disulfua (tức là protein multime). Ví dụ, polypeptit khảm bao gồm trình tự axit amin từ hai hoặc nhiều hơn hai nguồn có thể được xử lý thành dưới đơn vị, và các dưới đơn vị được kết hợp thông qua các cầu disulfua để tạo ra protein khảm hoặc polypeptit khảm.

Protein virut khảm theo các phương án khác nhau theo sáng chế bao gồm vùng xuyên màng và đuôi tế bào chất (TM/CT) từ HA bệnh cúm. TM/CT có thể là HA TM/CT tự nhiên từ loại, kiểu phụ bất kỳ thuộc virut bệnh cúm, gồm, ví dụ, loại hoặc kiểu phụ B, H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15 và H16. Các ví dụ không giới hạn của loại hoặc kiểu phụ H1, H3, H5 hoặc B gồm kiểu phụ A/NewCaledonia/20/99 (H1N1), kiểu phụ H1 A/California/04/09 (H1N1), kiểu phụ A/Indonesia/5/05 (H5N1), A/Brisbane/59/2007, B/Florida/4/2006, và H3 A/Brisbane/10/2007 (xem ví dụ WO2009/009876; WO 2009/076778; WO 2010/003225; WO 2010/003235, được kết hợp ở đây để tham khảo). Ngoài ra, TM/CT có thể loại bất kỳ trong số HA TM/CT trong đó 1, 2, 3, 4 hoặc 5 axit amin được loại bỏ, hoặc dựa trên loại miếng đệm bất kỳ hoặc trình tự nối liền kết của 1, 2, 3, 4 hoặc 5 axit amin được bổ sung vào.

Tốt hơn là, TM/CT là từ H5 hoặc H3, ví dụ nhưng không giới hạn với kiểu phụ A/Indonesia/5/05 (H5N1; “H5/Indo”; số truy cập ngân hàng gen ABW06108.1), H5 A/Vietnam/1194/2004(A-Vietnam; số truy cập ngân hàng gen ACR48874.1), H5 A/Anhui/1/2005 (A-Anhui; số truy cập ngân hàng gen ABD28180.1); H3 A/Brisbane/10/2007 (“H3/Bri”; số truy cập ngân hàng gen ACI26318.1), H3 A/Wisconsin/67/2005(A-WCN; số truy cập ngân hàng gen ABO37599.1). TM/CT của đường biên trong các trình tự H3 và H5 khác nhau được mô tả trong WO 2010/148511 (được kết hợp ở đây để tham khảo). Các ví dụ không giới hạn về trình tự axit amin đối với TM/CT gồm SEQ ID NO:41 và 42, và trình tự nucleotit bất kỳ mã hóa trình tự axit amin có SEQ ID NO:41 hoặc 42

Sự thay đổi axit amin được dung nạp trong hemagglutinin của virus cúm. Sự thay đổi này tạo ra các chủng mới liên tục được nhận biết. Tính lây nhiễm giữa các chủng mới có thể thay đổi. Tuy nhiên, sự tạo thành các trime hemagglutinin, sau đó tạo thành VLP được duy trì. Do đó, sáng chế đề xuất protein virus thể khảm bao gồm trình tự axit nucleic hemagglutinin, tạo thành VLP ở thực vật, và gồm các trình tự đã biết và trình tự HA biến thể có thể phát triển.

Thuật ngữ "hạt tương tự virus" (virus like particle - VLP), hoặc "các hạt tương tự virus" hoặc "VLP" dùng để chỉ các cấu trúc tự lắp ráp và bao gồm protein cấu trúc như protein virus thể khảm. VLP và VLP thể khảm thường về mặt hình thái học và tính kháng nguyên tương tự với các virion được tạo ra trong bệnh nhiễm trùng, nhưng thiếu thông tin di truyền đủ để tái bản và do đó không lây nhiễm. VLP và VLP thể khảm có thể được tạo ra trong tế bào chủ thích hợp bao gồm tế bào chủ ở thực vật. Sự chiết sau đây từ tế bào chủ và dựa trên sự phân lập và tinh chế thêm dưới các điều kiện thích hợp, VLP và VLP thể khảm có thể được tinh chế dưới dạng cấu trúc nguyên vẹn.

VLP thể khảm theo sáng chế có thể được tạo ra trong tế bào chủ được mô tả bằng sự thiếu khả năng sialyl hóa protein, ví dụ tế bào thực vật, tế bào sâu bọ, nấm, và các sinh vật khác bao gồm bọt biển, động vật ruột khoang, lớp giun đốt, động vật chân đốt, động vật thân mềm, lớp giun tròn, ngành trochelmintes, sán lá, hàm tơ, xúc tu, virus chlamydia, xoắn khuẩn, vi khuẩn gram dương, vi khuẩn lam, cổ vi khuẩn, hoặc loài tương tự. Xem, ví dụ Gupta và cộng sự, 1999. *Nucleic Acids Research* 27:370-372; Toukach và cộng sự, 2007. *Nucleic Acids Research* 35:D280-D286; Nakahara và cộng sự, 2008. *Nucleic Acids Research* 36:D368-D371.

Sáng chế cũng đề xuất VLP thể khảm thu được vỏ ngoài lipid từ màng tế bào của tế bào trong đó VLP thể khảm được biểu hiện. Ví dụ, nếu virus thể khảm được biểu hiện trong hệ thống trên cơ sở thực vật, VLP thể khảm tạo thành có thể thu được vỏ ngoài lipid từ màng tế bào của tế bào thực vật.

Nói chung, thuật ngữ "lipid" dùng để chỉ phân tử xuất hiện tự nhiên, hòa tan trong chất béo (ưa béo). VLP thể khảm được tạo ra ở thực vật theo một vài phương án của sáng chế có thể được tạo phức với lipid có nguồn gốc từ thực vật. Các lipid có nguồn gốc từ thực vật có thể có dạng lớp kép lipid, và có thể bao gồm thêm vỏ bao quanh VLP. Lipid có nguồn gốc từ thực vật có thể bao gồm các thành phần lipid của màng tế bào của thực vật trong đó VLP được tạo ra, bao gồm phospholipid, tri-, di- và

monoglycerit, cũng như sterol hòa tan trong chất béo hoặc chất chuyển hóa bao gồm sterol. Các ví dụ bao gồm phosphatidylcholin (phosphatidylcholine-PC), phosphatidyletanolamin (phosphatidylethanolamine-PE), phosphatidylinositol, phosphatidylserin, glycosphingolipit, phytosterol hoặc hỗn hợp của chúng. Lipit có nguồn gốc từ thực vật có thể lần lượt được đề cập đến 'lipit ở thực vật'. Các ví dụ về sterol thực vật bao gồm campesterol, stigmasterol, ergosterol, brasicasterol, denta-7-stigmasterol, denta-7-avenasterol, daunosterol, sitosterol, 24-metylcholesterol, cholesterol hoặc beta-sitosterol - xem, ví dụ, Mongrand và cộng sự, 2004. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này sẽ hiểu rằng, chế phẩm lipit trong màng tế bào của tế bào có thể thay đổi về canh trường và điều kiện sinh trưởng của tế bào hoặc sinh vật, hoặc loài, từ tế bào thu được. Nói chung, beta-sitosterol là sterol thực vật phong phú nhất.

Các màng tế bào thường bao gồm các lớp kép lipit, cũng như các protein có các chức năng khác nhau. Nồng độ hạn chế của lipit cụ thể có thể được tìm thấy trong cấu trúc lipit hai lớp, được đề cập đến dưới dạng 'số lượng lớn lipit'. Vùng rất nhỏ có lượng lớn lipit này có thể được làm giàu trong sphingolipit và sterol. Mà không ràng buộc bởi lý thuyết, các mảng lipit 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.

VLP được tạo ra ở thực vật có thể gây ra protein virus thể khảm bao gồm N-glycan đặc hiệu ở thực vật. Do đó, sáng chế cũng đề xuất VLP bao gồm protein virus thể khảm có N-glycan đặc hiệu ở thực vật.

Ngoài ra, sự cải biến N-glycan ở thực vật đã biết (xem ví dụ đơn xin cấp Patent Mỹ 60/944344; được kết hợp ở đây để tham khảo) và protein virus thể khảm có N-glycan được biến đổi có thể được tạo ra. Protein virus thể khảm bao gồm mẫu glycosyl hóa được biến đổi, ví dụ với N-glycan được fucosyl hóa, xylosyl hóa giảm, hoặc cả hai, được fucosyl hóa và xylosyl hóa có thể thu được, hoặc protein virus thể khảm có mẫu glycosyl hóa biến đổi có thể thu được, trong đó protein thiếu sự fucosyl hóa, xylosyl hóa, hoặc cả hai, và bao gồm sự galatosyl hóa tăng cao. Ngoài ra, sự cải biến của các cải biến sau dịch mã, ví dụ, bổ sung galactosa cuối cùng có thể gây ra sự giảm

fucosyl hóa và xylosyl hóa của protein virus thể khả khi được so sánh với thực vật loại đại biểu hiện protein virus thể khả.

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

Do đó sáng chế cũng đề xuất VLP bao gồm protein virus thể khả có N-glycan được biến đổi.

Không bị ràng buộc bởi lý thuyết, sự có mặt của N-glycan ở thực vật trên protein virus thể khả có thể kích thích đáp ứng miễn dịch bằng cách thúc đẩy liên kết của protein virus thể khả bằng các tế bào trình diện kháng nguyên. Sự kích thích đáp ứng miễn dịch sử dụng N glycan ở thực vật được đề xuất bởi Saint-Jore-Dupas và cộng sự (2007). Ngoài ra, hình dạng của VLP có thể là ưu điểm để trình diện kháng nguyên, và làm tăng tác dụng tá được của VLP khi được tạo phức với lớp lipid có nguồn gốc từ thực vật.

VLP có thể được đánh giá về cấu trúc và kích thước, ví dụ, bằng thử nghiệm hemagglutinin, kính hiển vi điện tử, hoặc phương pháp sắc ký loại trừ theo kích thước.

Đối với phương pháp sắc ký loại trừ theo kích thước, protein hòa tan tổng số có thể được chiết ra từ mô thực vật bằng cách làm đồng nhất (Polytron) mẫu vật liệu thực vật được nghiền đông lạnh trong đệm chiết, và vật liệu không hòa tan được loại bỏ bằng cách ly tâm. Sự kết tủa bằng axeton lạnh băng hoặc PEG cũng có thể có lợi. Protein hòa tan được định lượng, và dịch chiết được chuyển qua cột Sephacryl™, ví dụ cột Sephacryl™ S500. Blue Dextran 2000 có thể được sử dụng làm tiêu chuẩn định

chuẩn. Phương pháp sắc ký sau đây, các phân cắt có thể được phân tích bằng phương pháp thâm tách miễn dịch để xác định phân bổ sung protein vào phân cắt.

Phân tách riêng có thể ví dụ là phân nổi trên mặt (nếu được ly tâm, được lắng, hoặc kết tủa), hoặc phân lọc (nếu được lọc), và được làm giàu bằng protein, hoặc protein siêu cấu trúc, như ví dụ cấu trúc giống hình hoa thị hoặc bậc cao hơn, trọng lượng phân tử cao hơn, các hạt như VLP. Phần được tách ra có thể được xử lý để tách riêng, tinh chế, cô đặc hoặc hỗn hợp của chúng, protein, hoặc protein siêu cấu trúc, ví dụ, bằng các bước ly tâm bổ sung, bước sắc ký, kết tủa (ví dụ phương pháp sắc ký loại trừ theo kích thước, phương pháp sắc ký trao đổi ion, phương pháp sắc ký ái lực), lọc tiếp tuyến, hoặc kết hợp của chúng. Sự có mặt của protein được tinh chế, hoặc protein siêu cấu trúc, có thể được xác nhận, ví dụ, bằng phép phân tích Western, SDS-PAGE, hoặc tự nhiên sử dụng kháng thể phát hiện thích hợp, hiện tượng điện chuyển mao dẫn, kính hiển vi điện tử, hoặc phương pháp bất kỳ khác sẽ rõ ràng với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này.

Sơ đồ rửa giải có thể thay đổi phụ thuộc vào điều kiện rửa giải được sử dụng. Fig.7, Fig.8O và Fig.11B thể hiện ví dụ về sơ đồ rửa giải của phương pháp sắc ký loại trừ theo kích thước của dịch chiết thực vật chứa VLP thể khảm. Trong trường hợp này, VLP chứa HIV thể khảm, G bệnh dại thể khảm và protein bề mặt ba đơn phân của virus VZV rửa giải trong thể tích trống của cột, cấu trúc hình hoa thị và cấu trúc có trọng lượng phân tử cao rửa giải từ các phân tử 13 đến 14, và trọng lượng phân tử thấp hơn, hoặc dạng hòa tan của protein bề mặt ba đơn phân của virus thể khảm trong các phân tử 15 đến 17.

VLP có thể được tinh chế hoặc chiết bằng cách sử dụng phương pháp thích hợp bất kỳ ví dụ phương pháp chiết hóa học hoặc sinh hóa. VLP tương đối nhạy với việc sấy khô, nhiệt, độ pH, chất hoạt động bề mặt và chất tẩy. Do đó có thể hữu ích để sử dụng các phương pháp tối đa hóa hiệu suất, giảm thiểu sự nhiễm tạo của phân VLP do protein tế bào, duy trì tính nguyên vẹn của protein, hoặc VLP, và, cần đến, màng hoặc vỏ ngoài lipid được kết hợp, các phương pháp nối lỏng thành tế bào để giải phóng protein, hoặc VLP. Ví dụ, phương pháp tạo ra thể nguyên sinh và/hoặc tế bào có áo ngoài bị tiêu có thể được sử dụng (xem ví dụ WO 2011/035422, được kết hợp ở đây để tham khảo) thu được VLP như được mô tả ở trên. Giảm thiểu hoặc loại bỏ việc sử dụng chất tẩy rửa hoặc chất hoạt động bề mặt ví dụ SDS hoặc Triton X-100 có thể có

lợi để cải thiện hiệu suất chiết VLP. VLP tiếp đó có thể được đánh giá về cấu trúc và kích thước, ví dụ, bằng kính hiển vi điện tử, hoặc phương pháp loại trừ kích thước như được đề cập ở trên.

Một hoặc nhiều hơn một hoặc nhiều sản phẩm thiết kế di truyền thể khả theo sáng chế có thể được biểu hiện trong vật chủ thực vật thích hợp bất kỳ được biến nạp bằng trình tự nucleotit, hoặc sản phẩm thiết kế, hoặc vector theo sáng chế. Các ví dụ về vật chủ thích hợp gồm, nhưng không giới hạn với, cây nông nghiệp bao 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, hướng dương, bông và loại cây tương tự.

Một hoặc nhiều sản phẩm thiết kế di truyền thể khả theo sáng chế có thể còn bao gồm vùng không được dịch mã 3'. Vùng không được dịch mã 3' dùng để chỉ một phần của gen bao gồm đoạn ADN chứa tín hiệu polyadenyl hóa và tín hiệu điều hòa khác có khả năng ảnh hưởng đến việc xử lý mRNA hoặc sự biểu hiện gen. Tín hiệu polyadenyl hóa thường được mô tả bằng tác dụng bổ sung một lượng nhỏ axit polyadenylic vào đầu cuối 3' của tiền chất mRNA. Tín hiệu polyadenyl hóa nói chung được nhận diện bởi sự có mặt của tính tương đồng với dạng tiêu chuẩn tạo thành 5' AATAAA-3' mặc dù các biến thể là không phổ biến. Ví dụ không giới hạn về 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 (Ti) gen plasmit, như gen tổng hợp nopaline (nopaline synthase - NOS), gen thực vật như gen protein bảo quản đậu tương, dưới đơn vị nhỏ của gen ribuloza-I, 5-bisphosphat carboxylaza (ssRUBISCO; US 4962028; được kết hợp ở đây để tham khảo), trình tự khởi đầu được sử dụng để điều hòa sự biểu hiện plastocyanin, được mô tả trong US 7125978 (được kết hợp ở đây để tham khảo).

Một hoặc nhiều sản phẩm thiết kế di truyền thể khả theo sáng chế cũng có thể bao gồm vùng tăng cường, vùng tăng cường dịch mã hoặc phiên mã, có thể cần thiết. Vùng tăng cường có thể được định vị tại 5' hoặc 3' với trình tự được phiên mã. Vùng tăng cường đã biết bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, và có thể bao gồm codon khởi đầu ATG, trình tự liền kề hoặc trình tự tương tự. Codon khởi đầu, nếu có, có thể trong pha với khung đọc ("trong khung") của trình tự mã hóa tạo ra sự dịch mã chính xác của trình tự được phiên mã.

Sản phẩm thiết kế theo sáng chế có thể được đưa vào các tế bào của thực vật sử dụng Ti plasmit, Ri plasmit, vectơ virus ở thực vật, biến nạp ADN trực tiếp, vi tiêm, kỹ thuật xung điện, v.v.. Với các tổng kết về các kỹ thuật xem ví dụ Weissbach và Weissbach, *Methods for Plant Molecular Biology*, Academy Press, New York VIII, pp. 421-463 (1988); Geierman và Corey, *Plant Molecular Biology*, 2d Ed. (1988); và Miki và Iyer, *Fundamentals of Gene Transfer in Plants*. Trong *Plant Metabolism*, 2d Ed. DT. Dennis, DH Turpin, DD Lefebvre, DB Layzell (eds), Addison Wesley, Langmans Ltd. London, pp. 561-579 (1997). Các phương pháp khác bao gồm sự hấp thu ADN trực tiếp, sử dụng liposom, kỹ thuật xung điện, ví dụ sử dụng tế bào nguyên sinh, vi tiêm, bằng súng bắn gen hoặc sợi tinh thể, và sự thâm nhiễm chân không. Xem, ví dụ, Bilang, và cộng sự (Gen 100: 247-250 (1991), Scheid và cộng sự (Mol. Gen. Genet. 228: 104-112, 1991), Guerche và cộng sự (Plant Science 52: 111-116, 1987), Neuhauser và cộng sự (Theor. Appl Genet. 75: 30-36, 1987), Klein và cộng sự, Nature 327: 70-73 (1987); Howell và cộng sự (Science 208: 1265, 1980), Horsch và cộng sự (Science 227: 1229-1231, 1985), DeBlock và cộng sự, Plant Physiology 91: 694-701, 1989), *Methods for Plant Molecular Biology* (Weissbach và Weissbach, eds., Academic Press Inc., 1988), *Methods in Plant Molecular Biology* (Schuler và Zielinski, eds., Academic Press Inc., 1989), Liu và Lomonosoff (J Virol Meth, 105:343-348, 2002), các Patent Mỹ Số 4945050; 5036006; và 5100792, các đơn xin cấp Patent Mỹ số 08/438666, nộp ngày 10 tháng 05, 1995, và 07/951715, nộp ngày 25 tháng 9, 1992, (tất cả được kết hợp ở đây để tham khảo).

Như được mô tả dưới đây, phương pháp biểu hiện tạm thời có thể được sử dụng để biểu hiện sản phẩm thiết kế theo sáng chế (xem Liu và Lomonosoff, 2002, *Journal of Virological Methods*, 105:343-348; được kết hợp ở đây để tham khảo). Mặc khác, phương pháp biểu hiện tạm thời trên cơ sở chân không, như được mô tả bởi Kapila và cộng sự, 1997, được kết hợp ở đây để tham khảo) có thể được sử dụng. Các phương pháp này có thể bao gồm, ví dụ, nhưng không giới hạn với, phương pháp tiêm Agro hoặc thâm nhiễm Agro, thâm nhiễm qua ống tiêm, tuy nhiên, phương pháp tạm thời khác cũng có thể được sử dụng như được chỉ rõ ở trên. Với sự tiêm Agro, sự thâm nhiễm Agro, hoặc sự thâm nhiễm qua ống tiêm, hỗn hợp của *Agrobacteria* bao gồm axit nucleic mong muốn đi vào không gian nội bào của mô, ví dụ lá, phần trên không của thực vật (gồm thân, lá và hoa), phần khác của thực vật (thân, rễ, hoa), hoặc toàn bộ

thực vật. Sau khi lai giống biểu bì *Agrobacteria* nhiễm độc và truyền bản sao t-ADN vào tế bào. t-ADN được phiên mã thuộc thể bổ sung và mARN được dịch mã, dẫn đến việc sản xuất protein được quan tâm trong các tế bào được tiêm nhiễm, tuy nhiên, sự chuyển t-ADN bên trong nhân chỉ là tạm thời.

Để hỗ trợ việc nhận biết các tế bào thực vật được biến nạp, các sản phẩm thiết kế theo sáng chế có thể được thao tác để gồm các chỉ thị có thể chọn được của thực vật. Các chỉ thị có thể chọn được hữu ích gồm các enzym tạo ra tính bền với các chất hóa học như thuốc kháng sinh ví dụ, gentamycin, hygromycin, kanamycin, hoặc thuốc diệt cỏ như phosphinothryxin, glyphosat, clorosulfuron, và các chất tương tự. Tương tự, các enzym tạo ra việc sản xuất hợp chất có thể được nhận biết bằng sự thay đổi màu sắc như GUS (beta-glucuronidaza), hoặc sự phát quang, như luciferaza hoặc GFP, có thể được sử dụng.

Ngoài ra, phần được xem xét theo sáng chế là các thực vật chuyển gen, các tế bào hoặc hạt thực vật chứa sản phẩm thiết kế gen thể khả năng theo sáng chế. Các phương pháp tạo ra toàn bộ cây từ các tế bào thực vật cũng đã được biết trong tình trạng kỹ thuật. Nói chung, các tế bào thực vật được 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 chất chỉ thị có thể chọn được để nhận biết dễ dàng các tế bào thực vật được biến nạp. Một lần nữa tạo thành mô sẹo, sự tạo thành chồi có thể được hỗ trợ nhờ sử dụng các hormon thích hợp của thực vật 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 thực vật chuyển gen cũng có thể được tạo ra mà không sử dụng nuôi cấy mô.

Sáng chế bao gồm các trình tự nucleotit:

Bảng 1. Danh sách các số nhận biết trình tự.

SEQ ID NO:	Phần mô tả	Bảng/Fig.
1	Trình tự axit nucleic liên ứng của HIV ConS ΔCFI (Peptit tín hiệu tự nhiên được in đậm, vùng xuyên màng tự nhiên và vùng chất nguyên sinh được gạch dưới)	Fig.1A
2	IF-ApaI-SpPDI.c	Fig.1B
3	SpPDI-HIV gp145.r	Fig.1C

4	IF-SpPDI-gp145.c	Fig.1D
5	WtdTm-gp145.r	Fig.1E
6	Catxet biểu hiện số 995, từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc NOS). Vị trí SbfI được loại trừ được viết đậm. PDISP-HIV ConS ΔCFI được gạch chân.	Fig.1F
7	Trình tự axit amin của PDISP-HIV ConS ΔCFI	Fig.1G
8	IF-H3dTm+gp145.r	Fig.2A
9	Gp145+H3dTm.c	Fig.2B
10	H3dTm.r	Fig.2C
11	Catxet biểu hiện số 997, từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc NOS). Vị trí SbfI được loại trừ được viết đậm. PDISP-HIV Con S ΔCFI-A/Brisbane/10/2007 H3 TM+CT được gạch dưới.	Fig.2D
12	Trình tự axit amin của PDISP-HIV ConS ΔCFI-A/Brisbane/10/2007 H3 TM+CT	Fig.2E
13	IF-H5dTm+gp145.r	Fig.3A
14	Gp145+H5dTm.c	Fig.3B
15	IF-H5dTm.r	Fig.3C
16	Catxet biểu hiện số 999, từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc NOS). Vị trí SbfI được loại trừ được viết đậm. PDISP-HIV ConS ΔCFI-A/Indonesia/5/2005 H5 TM+CT được gạch dưới.	Fig.3D
17	Trình tự axit amin của PDISP-HIV ConS ΔCFI-A/Indonesia/5/2005 H5 TM+CT	Fig.3E
18	Catxet biểu hiện số 172, từ XmaI (phía trước trình tự khởi đầu plastoxyanin) đến EcoRI (ngay sau trình tự kết thúc plastoxyanin). Trình tự axit nucleic TBSV P19 được gạch dưới.	Fig.4A
19	Trình tự axit amin của gen ức chế TBSV P19 bị bất hoạt	Fig.4B
20	IF-wtSp-VZVgE.c	Fig.10A
21	IF-H5dTm+VZVgE.r	Fig.10B
22	Gen VZV gE được tổng hợp (tương ứng với nt 3477-5348 từ số truy cập ngân hàng gen AY013752.1)	Fig.10C

23	VZVgE+H5dTm.c	Fig.10D
24	Catxet biểu hiện số 946, từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc NOS).VZV gE-A/Indonesia/5/2005 H5 TM+CT được gạch dưới	Fig.10E
25	Trình tự axit amin của VZV gE- A/Indonesia/5/2005 H5 TM+CT	Fig.10F
26	IF-wtSp-SARSgS.c	Fig.12B
27	IF-H5dTm+SARSgS.r	Fig.12C
28	gen SARS gS được tổng hợp (tương ứng với nt 21492-25259 từ số truy cập ngân hàng gen AY278741.1) (Peptit tín hiệu tự nhiên được in đậm, vùng xuyên màng tự nhiên và vùng chất nguyên sinh được gạch dưới)	Fig.12D
29	SARSgS+H5dTm.c	Fig.12E
30	Catxet biểu hiện số 916, từ PacI (phía trước trình tự khởi đầu) đến AscI (ngay sau trình tự kết thúc NOS). SARS gS-A/Indonesia/5/2005 H5 TM+CT được gạch dưới	Fig.12F
31	Trình tự axit amin của SARS gS- A/Indonesia/5/2005 H5 TM+CT	Fig.12G
32	IF-RabG-S2+4.c	Fig.8B
33	RabG+H5dTm.r	Fig.8C
34	Gen Rab G được tổng hợp (tương ứng với nt 3317-4891 từ số truy cập ngân hàng gen EF206707) (Peptit tín hiệu tự nhiên được in đậm, vùng xuyên màng tự nhiên và vùng chất nguyên sinh được gạch dưới)	Fig.8D
35	IF-H5dTm.c	Fig.8E
36	Sản phẩm thiết kế 141 t-ADN từ trái sang phải (được gạch dưới). Hệ thống biểu hiện 2X35S-CPMV-HT-PDISP-NOS với catxet biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin	Fig.8F
37	Catxet biểu hiện số 1074 từ trình tự khởi đầu 2X35S đến trình tự kết thúc NOS. PDISP-Rab G -A/Indonesia/5/2005 H5 TM+CT được gạch dưới.	Fig.8H
38	Trình tự axit amin của PDISP-Rab G- A/Indonesia/5/2005 H5 TM+CT	Fig.8I
39	Sản phẩm thiết kế 144 t-ADN từ trái sang phải (được gạch dưới). 2X35S-CPMV-HT-PDISP-NOS thành BeYDV +	Fig.8J

	hệ thống biểu hiện tái bản với catxet biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin	
40	Catxet biểu hiện số 1094 BeYDV LIR.PDISP-Rab G - A/Indonesia/5/2005 H5 TM+CT từ phải sang trái được gạch dưới.	Fig.8L
41	H5 (A/Indonesia/05/2005) TM/CT: QILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRI CI	
42	H3 (A/Brisbane/10/2007) TM/CT: DWILWISFAISCFLLCVALLGFIMWACQKGNIRCNIC I	
43	IF-Opt_EboGP.s2+4c	Fig.13A
44	H5iTMCT+Opt_EboGP.r	Fig.13B
45	Gen GP tổng hợp được tối ưu hóa	Fig.13C
46	Opt_EboGP+H5iTMCT.c	Fig.13D
47	Sản phẩm thiết kế 1192	Fig.13F
48	Catxet biểu hiện số 1366	Fig.13G
49	Trình tự axit amin của PDISP-GP từ Zaire 95 Ebolavirus-H5 TM+CT từ A/Indonesia/5/2005	Fig.13 H

Sáng chế sẽ được minh họa thêm trong các ví dụ dưới đây.

Ví dụ thực hiện sáng chế

Sản phẩm thiết kế

Bảng 2: Sản phẩm thiết kế bao gồm các trình tự mã hóa protein HIV

Hệ thống biểu hiện	Hệ thống phóng đại	Sp	5'	Protein	3'	NptII (Ori tADN)	Plasto-P19 (Ori tADN)	Sản phẩm thiết kế số
CPMV HT	-	Wt Sp	-	Gp145ΔCFI	wt TM+CT	L→R	-	994
CPMV HT	-	Sp PDI	-	Gp145ΔCFI	wt TM+CT	L→R	-	995
CPMV HT	-	Wt Sp	-	Gp145ΔCFI	H3 A/Bri TM+CT	L→R	-	996
CPMV HT	-	Sp PDI	-	Gp145ΔCFI	H3 A/Bri TM+CT	L→R	-	997
CPMV HT	-	Wt Sp	-	Gp145ΔCFI	H5 A/Indo TM+CT	L→R	-	998
CPMV HT	-	Sp PDI	-	Gp145ΔCFI	H5 A/Indo TM+CT	L→R	-	999
CPMV HT	BeYDV+rep	Sp PDI	-	Gp145ΔCFI	wt TM+CT	-	L→R	985
CPMV HT	BeYDV+rep	Sp PDI	-	Gp145ΔCFI	H3 A/Bri TM+CT	-	L→R	987
CPMV HT	BeYDV+rep	Sp PDI	-	Gp145ΔCFI	H5 A/Indo TM+CT	-	L→R	989

Bảng 3: Sản phẩm thiết kế bao gồm các trình tự mã hóa protein virut gây bệnh dại

Hệ thống biểu hiện	Hệ thống phóng đại	Sp	5'	Protein	3'	NptII (Ori tDNA)	Plasto-P19 (Ori tADN)	Sản phẩm thiết kế số
CPMV HT	-	wt SP	-	G bệnh dại	-	-	L→R	1070
CPMV HT	-	Sp PDI	-	G bệnh dại	-	-	L→R	1071
CPMV HT	BeYDV+rep	wt SP	-	G bệnh dại	-	-	L→R	1090
CPMV HT	BeYDV+rep	Sp PDI	-	G bệnh dại	-	-	L→R	1091
CPMV HT	-	Sp PDI	-	G bệnh dại	H5 A/Indo TM+CT	-	L→R	1074
CPMV HT	BeYDV+rep	Sp PDI	-	G bệnh dại	H5 A/Indo TM+CT	-	L→R	1094
CPMV HT	BeYDV+rep	Sp PDI	-	G bệnh dại (A447S)	-	-	L→R	1072
CPMV HT	BeYDV+rep	Sp PDI	-	G bệnh dại (A447S)	-	-	L→R	1092
CPMV HT	-	Sp PDI	-	G bệnh dại (M44I+V392G)	-	-	L→R	1073
CPMV HT	BeYDV+rep	Sp PDI	-	G bệnh dại (M44I+V392G)	-	-	L→R	1093
CPMVHT		Sp PDI	-	G bệnh dại (M44I+V392G)	H5 A/Indo TM+CT	-	L→R	1075
CPMVHT	BeYDV+rep	Sp PDI	-	G bệnh dại (M44I+V392G)	H5 A/Indo TM+CT	-	L→R	1095

Bảng 4: Sản phẩm thiết kế bao gồm các trình tự mã hóa protein VZV

Hệ thống biểu hiện	Hệ thống phóng đại	Sp	5'	Protein	3'	NptII (Ori Tdna)	Plasto-P19 (Ori tADN)	Sản phẩm thiết kế số
CPMV HT	-	Wt Sp	-	VZV gE	wt TM+CT	L→R		944
CPMV HT	-	Sp PDI	-	VZV gE	wt TM+CT	L→R	-	945
CPMV HT	-	Wt Sp	-	VZV gE	HS A/Indo TM+CT	L→R	-	946
CPMV HT	-	Sp PDI	-	VZV gE	HS A/Indo TM+CT	L→R	-	947

Bảng 5: Sản phẩm thiết kế bao gồm các trình tự mã hóa protein SARS

Hệ thống biểu hiện	Hệ thống phóng đại	Sp	5'	Protein	3'	NptII (Ori tADN)	Plasto-P19 (OritADN)	Sản phẩm thiết kế số
CPMV HT		Wt Sp	-	SARS gS	wt TM+CT	L→R		914
CPMV HT	-	Sp PDI	-	SARS gS	wt TM+CT	L→R	-	915
CPMV HT		Wt Sp	-	SARS gS	H5 A/Indo TM+CT	L→R	-	916
CPMV HT		Sp PDI	-	SARS gS	H5 A/Indo TM+CT	L→R	-	917

Ví dụ 1: Lắp ráp catxet biểu hiện với protein HIV

2X35S-CPMV HT –PDISP-HIV ConS Δ CFI-NOS (sản phẩm thiết kế số 995)

Trình tự mã hóa peptit tín hiệu PDI ở cở linh lăng được dung hợp vào HIV ConS Δ CFI bằng vùng xuyên màng tự nhiên và đuôi chất nguyên sinh được tách dòng thành hệ thống biểu hiện 2X35S-CPMV-HT như sau. Đầu tiên, trình tự axit nucleic của gen HIV ConS Δ CFI, bao gồm peptit tín hiệu tự nhiên và vùng xuyên màng và vùng tế bào chất được tổng hợp bởi GeneArt AG (Regensburg, Germany) theo các trình tự được bộc lộ trong Liao và cộng sự, (2006, Virology 353: 268-282). Trình tự axit nucleic của HIV ConS Δ CFI được thể hiện trên Fig.1A (SEQ ID NO: 1). Peptit tín hiệu của protein disulfua isomeraza ở cở linh lăng (protein disulfide isomerase - PDISP) (nucleotit 32-103; số truy cập Z11499) được liên kết với HIV ConS Δ CFI bằng phương pháp gắn trên cơ sở PCR được thể hiện bởi Darveau và cộng sự (Methods in Neuroscience 26: 77-85 (1995)). Trong chu kỳ PCR thứ nhất, đoạn chứa PDISP được phóng đại bằng cách sử dụng môi IF-ApaI-SpPDI.c (Fig.1B, SEQ ID NO: 2) và SpPDI-HIV gp145.r (Fig.1C, SEQ ID NO: 3) với sản phẩm thiết kế số 540 (xem Fig.52, SEQ ID NO: 61 trong WO 2010/003225 mà được kết hợp ở đây để tham khảo, với trình tự của sản phẩm thiết kế số 540) làm khuôn. Đoạn thứ hai chứa ConS Δ CFI không phải peptit tín hiệu tự nhiên được phóng đại bằng cách sử dụng môi IF-SpPDI-gp145.c (Fig.1D, SEQ ID NO: 4) và WtdTm-gp145.r (Fig.1E, SEQ ID NO: 5) sử dụng đoạn ConS Δ CFI được tổng hợp (Fig.1A, SEQ ID NO: 1) làm khuôn. Trong chu kỳ PCR thứ hai, các môi IF-ApaI-SpPDI.c (Fig.1B, SEQ ID NO: 2) và WtdTm-gp145.r (Fig.1E, SEQ ID NO: 5) được sử dụng cùng với cả hai sản phẩm PCR từ chu kỳ PCR thứ nhất làm khuôn. Sản phẩm PCR tạo thành được tiêu hóa bằng enzym giới hạn ApaI và được tách dòng thành sản phẩm thiết kế 972 được biến đổi được tiêu hóa bằng ApaI-StuI. Plasmid chất nhận 972 được biến đổi (xem Fig.94, SEQ ID NO: 134 trong WO 2010/003225, được kết hợp ở đây để tham khảo, với trình tự ban đầu của sản phẩm thiết kế số 972) được xử lý để loại bỏ vị trí giới hạn SbfI phía trước trình tự khởi đầu 2X35S. Vị trí SbfI được loại bỏ bằng cách tiêu hóa plasmid 972 bằng SbfI, xử lý plasmid tạo thành bằng T4 ADN polymeraza để loại bỏ 3' nhô ra ở trên và gắn lại plasmid được xử lý, dẫn đến plasmid 972 mà không có vị trí SbfI. Sản phẩm thiết kế tạo thành là số 995 đã biết (Fig.1F, SEQ ID NO: 6). Trình tự axit amin có PDISP-HIV

ConS Δ CFI được thể hiện trên Fig.1G (SEQ ID NO: 7). Sự thể hiện plasmit 995 được thể hiện trên Fig.1H. Sản phẩm thiết kế này không mã hóa protein M1.

2X35S-CPMV HT –PDISP-HIV ConS Δ CFI+ H3 A/Brisbane/10/2007 (TmD + Cyto tail)-NOS (sản phẩm thiết kế số 997)

Trình tự mã hóa peptit tín hiệu PDI ở cở linh lăng được dung hợp với HIV ConS Δ CFI và với vùng xuyên màng và vùng chất nguyên sinh của H3 A/Brisbane/10/2007 được tách dòng thành hệ thống biểu hiện 2X35S-CPMV-HT sử dụng phương pháp gắn trên cơ sở PCR được thể hiện bởi Darveau và cộng sự (Methods in Neuroscience 26: 77-85 (1995)). Trong chu kỳ PCR thứ nhất, đoạn bao gồm PDISP-HIV ConS Δ CFI mà không có vùng xuyên màng tự nhiên và vùng tế bào chất được phóng đại bằng cách sử dụng môi IF-ApaI-SpPDI.c (Fig.1B, SEQ ID NO: 2) và IF-H3dTm+gp145.r (Fig.2A, SEQ ID NO: 8), sử dụng sản phẩm thiết kế số 995 (Fig.1F, SEQ ID NO: 6) làm khuôn. Đoạn thứ hai chứa vùng xuyên màng và vùng chất nguyên sinh của H3 A/Brisbane/10/2007 được phóng đại bằng cách sử dụng môi Gp145+H3dTm.c (Fig.2B, SEQ ID NO: 9) và H3dTm.r (Fig.2C, SEQ ID NO: 10), sử dụng sản phẩm thiết kế số 776 (xem Fig.60, SEQ ID NO: 69 trong WO 2010/003225, được kết hợp ở đây để tham khảo, với trình tự sản phẩm thiết kế số 776) làm khuôn. Sản phẩm PCR từ cả hai đoạn phóng đại tiếp đó được trộn và được sử dụng làm khuôn cho chu kỳ phóng đại thứ hai sử dụng IF-ApaI-SpPDI.c (Fig.1B, SEQ ID NO: 2) và H3dTm.r (Fig.2C, SEQ ID NO: 10) làm môi. Sản phẩm của phản ứng PCR thứ hai tiếp đó được tiêu hóa bằng ApaI và được tách dòng thành sản phẩm thiết kế số 995 (Fig.1F, SEQ ID NO: 6) được phân hủy bằng enzym giới hạn ApaI và StuI. Sản phẩm thiết kế tạo thành là số 997 đã biết (Fig.2D, SEQ ID NO: 11). Trình tự axit amin của PDISP-HIV ConS Δ CFI-A/Brisbane/10/2007 H3 TM+CT được thể hiện trên Fig.2E (SEQ ID NO: 12). Sự thể hiện plasmit 997 được thể hiện trên Fig.2F. Sản phẩm thiết kế này không mã hóa protein M1.

2X35S-CPMV HT–PDISP-HIV ConS Δ CFI-H5 A/Indonesia/5/2005 (TmD + Cyto tail)-NOS (sản phẩm thiết kế số 999)

Trình tự mã hóa peptit tín hiệu PDI ở cở linh lăng được dung hợp với HIV ConS Δ CFI và với vùng xuyên màng và vùng chất nguyên sinh của H5 A/Indonesia/5/2005 được tách dòng thành hệ thống biểu hiện 2X35S-CPMV-HT như sau sử dụng phương pháp gắn trên cơ sở PCR được thể hiện bởi Darveau và cộng sự (Methods in

Neuroscience 26: 77-85 (1995)). Trong chu kỳ PCR thứ nhất, đoạn chứa PDISP-HIV ConS Δ CFI mà không có vùng xuyên màng tự nhiên và vùng tế bào chất được phóng đại bằng cách sử dụng mồi IF-ApaI-SpPDI.c (Fig.1B, SEQ ID NO: 2) và IF-H5dTm+gp145.r (Fig.3A, SEQ ID NO: 13), sử dụng sản phẩm thiết kế số 995 (Fig.1F, SEQ ID NO: 6) làm khuôn. Đoạn thứ hai chứa vùng xuyên màng và vùng tế bào chất của H5 A/Indonesia/5/2005 được phóng đại bằng cách sử dụng mồi Gp145+H5dTm.c (Fig.3B, SEQ ID NO: 15) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15), sử dụng sản phẩm thiết kế số 972 (xem Fig.94, SEQ ID NO: 134 trong WO 2010/003225, được kết hợp ở đây để tham khảo, với trình tự của sản phẩm thiết kế số 972) làm khuôn. Sản phẩm PCR từ cả hai đoạn phóng đại tiếp đó được trộn và được sử dụng làm khuôn cho chu kỳ phóng đại thứ hai sử dụng IF-ApaI-SpPDI.c (Fig.1B, SEQ ID NO: 2) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15) làm mồi. Sản phẩm của phản ứng PCR thứ hai tiếp đó được tiêu hóa bằng ApaI và được tách dòng thành sản phẩm thiết kế số 995 (Fig.1F, SEQ ID NO: 6) được tiêu hóa bằng enzym giới hạn ApaI và StuI. Sản phẩm thiết kế tạo thành là số 999 đã biết (Fig.3D, SEQ ID NO: 16). Trình tự axit amin của PDISP-HIV ConS Δ CFI- A/Indonesia/5/2005 H5 TM+CT. TM+CT được thể hiện trên Fig.3E (SEQ ID NO: 17). Sự thể hiện plasmit 999 được thể hiện trên Fig.3F. Sản phẩm thiết kế này không mã hóa protein M1.

Plastoxyanin-P19-plastoxyanin (sản phẩm thiết kế số 172)

Trình tự mã hóa gen ức chế P19 để bất hoạt từ virus gây còi cọc ở bụi cây cà chua (Tomato Bushy Stunt Virus - TBSV) được tách dòng giữa trình tự khởi đầu plastoxyanin cở linh lăng và 3'UTR và trình tự kết thúc như sau. Sản phẩm thiết kế số R472 (xem WO 2010/003225 A1, được kết hợp ở đây để tham khảo, để lắp ráp và Fig.86 trong WO 2010/003225 A1 để thể hiện plasmit R472) được tiêu hòa bằng enzym giới hạn DraIII (84 cặp bazơ phía trước ATG khởi đầu) và SacI (9 cặp bazơ phía sau codon kết thúc) để loại bỏ đoạn bao gồm 84 pb từ trình tự khởi đầu plastoxyanin ở cở linh lăng và trình tự mã hóa gen ức chế TBSV P19 để bất hoạt. Đoạn tạo thành được tách dòng thành sản phẩm thiết kế 540 (xem WO 2010/003225, được kết hợp ở đây để tham khảo, để lắp ráp và Fig.6 của Patent này với sự thể hiện plasmit 540) được tiêu hóa trước bằng DraIII và SacI. Sản phẩm thiết kế tạo thành là số 172 đã biết (SEQ ID NO: 4A, Fig.18). Trình tự axit amin của protein TBSV P19

được thể hiện trên Fig.4B (SEQ ID NO: 19). Sự thể hiện plasmid 172 được thể hiện trên Fig.4C.

Chế phẩm sinh khối của thực vật, chất để tiêm và sự thâm nhiễm *agrobacterium*

Thuật ngữ "sinh khối" và "chất của thực vật" được sử dụng ở đây được hiểu là để phản ánh vật liệu bất kỳ có nguồn gốc từ thực vật. Sinh khối hoặc chất của thực vật có thể chứa toàn bộ thực vật, mô, tế bào, hoặc phần bất kỳ khác. Ngoài ra, sinh khối hoặc chất của thực vật có thể bao gồm thành phần nội bào của thực vật, thành phần ngoại bào của thực vật, dịch chiết lỏng hoặc rắn của thực vật, hoặc hỗn hợp của chúng. Ngoài ra, sinh khối hoặc chất của thực vật có thể bao gồm thực vật, tế bào của thực vật, mô, dịch chiết lỏng, hoặc hỗn hợp của chúng, từ lá, thân, quả, rễ hoặc hỗn hợp của chúng. Một bộ phận của thực vật có thể bao gồm chất của thực vật hoặc sinh khối.

Cây *Nicotiana benthamiana* được phát triển từ hạt giống trong vùng đất phẳng chứa đầy chất 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, nhỏ 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.

Agrobacteria được gây nhiễm bằng mỗi sản phẩm thiết kế được phát triển trong môi trường YEB có bổ sung 10mM axit 2-(N-morpholino)etansulfonic (MES), 20μM axetosyringon, 50μg/ml kanamycin và 25μg/ml carbenicillin 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 (10mM MgCl₂ và 10mM MES pH=5,6). và được bảo quản qua đêm ở nhiệt độ 4°C. Vào ngày thâm nhiễm, các mẻ canh trường được pha loãng trong 2,5 thể tích canh trường và được để ấm trước khi sử dụng. Toàn bộ cây *N. benthamiana* được đặt lộn ngược xuống trong huyền phù vi khuẩn trong thùng thép không gỉ kín không khí dưới chân không có áp suất 20-40 Torr trong thời gian 2 phút. Cây được đưa trở lại nhà kính trong thời gian từ 2 đến 6 ngày ủ cho đến khi thu hoạch. Trừ khi có quy định khác, tất cả các lần thâm nhiễm được tiến hành là thâm nhiễm đồng thời bằng chủng AGL1/172 với tỷ lệ 1:1.

Chủng *A. tumefaciens* bao gồm sản phẩm thiết kế khác nhau được mô tả ở đây được đề cập đến việc sử dụng tiền tố "AGL1". Ví dụ *A. tumefaciens* bao gồm sản phẩm thiết kế số 995 (xem Fig.1H, được gọi là "AGL1/995").

Thu hoạch lá và dịch chiết protein tổng số

Sau khi ủ, phần trên không của cây được thu hoạch, được làm đông lạnh ở nhiệt độ -80°C và nghiền thành các mẫu. Protein hòa tan tổng số được chiết bằng cách làm đồng nhất (Polytron) mỗi mẫu vật liệu thực vật được nghiền lạnh trong 3 thể tích 50mM Tris pH=8,0 lạnh, 0,15M NaCl, 0,1% Triton X-100 và 1mM phenylmetansulfonyl florua. Sau khi đồng nhất hóa, huyền phù đặc được ly tâm ở 10000g trong thời gian 10 phút ở nhiệt độ 4°C và dịch chiết thô được lọc này (phần nổi trên mặt) được giữ lại để phân tích.

Phân tích protein và thử nghiệm thẩm tách miễn dịch

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. Protein được tách ra bằng SDS-PAGE và được chuyển điện tử lên trên màng polyvinyliden diflorua (PVDF) (Roche Diagnostics Corporation, Indianapolis, IN) để phát hiện miễn dịch. Trước khi thẩm 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 .

Phương pháp thẩm tách miễn dịch được tiến hành bằng cách ủ bằng kháng thể thứ cấp kháng gp120 đa dòng ở dê (Abcam, ab21179) được pha loãng 1:2500 trong $2\mu\text{g/ml}$ trong 2% sữa không kem trong TBS-Tween 20 0,1%. Phát hiện sự phát quang bằng phản ứng hóa học được tiến hành sau khi ủ với kháng thể thứ cấp được tiếp hợp peroxidaza (JIR 705-035-147), được pha loãng 1:10000 trong 2% sữa không kem trong TBS-Tween 20 0,1%. Các phức hợp phản ứng miễn dịch được phát hiện bằng sự phát quang bằng phản ứng hóa học sử dụng luminol làm cơ chất (Roche Diagnostics Corporation).

Ví dụ 2: Sự biểu hiện của protein vỏ ngoài HIV tự nhiên và thể khảm ở thực vật

HIV ConS ΔCFI là protein vỏ ngoài M nhóm HIV liên ứng với các vòng biến thiên được rút ngắn, sự phát hiện vị trí cắt, vùng dung hợp và vùng chiếm ưu thế miễn dịch trong gp41. Điều này được chứng minh để tạo ra các kháng thể trung hòa kiểu phụ chéo có bề rộng tương tự hoặc lớn hơn và chuẩn hơn so với protein vỏ ngoài loại đại ở chuột lang (Liao và cộng sự, Virology (2006) 353: 268-282).

Fig.5 thể hiện sản phẩm thiết kế được sử dụng trong nghiên cứu này. Trong sản phẩm thiết kế số 995, vùng mã hóa với HIV ConS ΔCFI trưởng thành (sau đây được

đề cập đến dưới dạng Env), bao gồm TM/CT tự nhiên, được đặt dưới sự kiểm soát của hệ thống biểu hiện CPMV-HT. Sản phẩm thiết kế số 997 và 999, vùng TM/CT của HIV Env được thay thế bằng vùng hemagglutinin cúm (hemagglutinin - HA) từ A/Brisbane/10/2007 (H3N2) và A/Indonesia/05/2005 (H5N1), tương ứng. Trong tất cả các trường hợp, peptit tín hiệu có nguồn gốc từ thực vật – từ protein disulfua isomeraza ở cỏ linh lăng (protein disulfide isomerase - PDI) – được thay thế peptit tín hiệu protein HIV Env tự nhiên.

Việc sản xuất HIV Env từ các sản phẩm thiết kế 995, 997 và 999 được so sánh trong các cây *Nicotiana benthamiana* được thâm nhiễm agrobacterium. Dịch chiết protein từ cây được thâm nhiễm bằng AGL1/995, AGL1/997 và AGL1/999 được phân tích bằng phương pháp thẩm tách Western, và kết quả được thể hiện trên Fig.6, trong đó các làn từ 1 đến 4 là đối chứng dương chứa các lượng HIV-1 gp160 (ab68171) tái tổ hợp khác nhau; làn 5, đối chứng âm; làn 6 đến 8, protein được chiết từ các lá được thâm nhiễm AGL1/995; làn 9 và 10, protein được chiết từ lá được thâm nhiễm AGL1/997; làn 11 và 12, protein được chiết từ các lá được thâm nhiễm AGL1/999.

Như được thể hiện trên Fig.6, sự biểu hiện của HIV Env tự nhiên không thể phát hiện được trong các điều kiện được sử dụng để phát hiện, xác nhận mức tích tụ rất thấp trước đây được thông báo đối với protein HIV Env ở thực vật (Rybicki và cộng sự, 2010, Plant Biotechnology Journal 8: 620-637). Dạng khảm của Env, trong đó các vùng TM/CT được thay thế bởi vùng H3 bệnh cúm (sản phẩm thiết kế # 997) hoặc H5 (sản phẩm thiết kế #999) được tích tụ ở mức cao hơn nhiều so với dạng tự nhiên. Như được lưu ý ở trên, sản phẩm thiết kế #997 và #999 không mã hóa protein M1.

Ví dụ 3: HIV Env thể khảm lắp ráp thành các hạt tương tự virus

Khả năng của HIV Env thể khảm để có thể lắp ráp thành VLP ở cây khi không có mặt của protein lõi và protein nền cũng được nghiên cứu. Protein chiết từ AGL1/997 (Env/H3) và AGL1/999 (Env/H5) phụ thuộc vào phương pháp sắc ký lọc gel sau đó bằng cách phân tích các phần rửa giải đối với hàm lượng Env sử dụng các kháng thể kháng gp120 ở dê. Kết quả thẩm tách Western được thể hiện trên Fig.7A chỉ ra rằng đối với dịch chiết từ cây được thâm nhiễm AGL1/999, hàm lượng Env/H5 thể khảm trong các phần rửa giải cực đại trong các phần từ 7 đến 10, chỉ ra sự lắp ráp chúng trong các cấu trúc có trọng lượng phân tử rất cao lớn hơn 2 triệu dalton. Việc nghiên cứu hàm lượng protein tương đối trong các phần từ 7 đến 18 rõ ràng chỉ ra rằng

phần lớn protein ở vật chủ trong các phần từ 11 đến 18 (Fig.7B). Các kết quả này chỉ ra rằng Env/H5 lắp ráp thành cấu trúc có trọng lượng phân tử cao có kích thước nằm ngoài trọng lượng phân tử mong muốn của tam phân đồng nhất. Các kết quả tương tự thu được đối với Env/H3 trong cây được thâm nhiễm AGL1/997. Các kết quả này được đưa ra cùng nhau chứng minh rằng HIV Env/H5 thể khảm được tích tụ ở mức cao trong các cây được thâm nhiễm agrobacterium và được lắp ráp thành các hạt tương tự virus khi không có mặt protein lõi hoặc protein nền. Các kết quả này cũng chứng minh rằng sự dung hợp của vùng TM/CT của HA cúm thành các kháng nguyên không phải bệnh cúm đủ để lắp ráp và giải phóng các VLP trình diện kháng nguyên không phải bệnh cúm.

Ví dụ 4: Sự lắp ráp catxet biểu hiện bằng protein bệnh dại

Glycoprotein G bệnh dại C-2X35S-CPMV HT-PDISP (RabG)+ vùng xuyên màng và đuôi tế bào chất H5 A/Indonesia/5/2005 (TM+CT)-NOS (Sản phẩm thiết kế số 1074; Fig.8A, Fig.8H).

Trình tự mã hóa vùng ngoại bào (RabG) glycoprotein G bệnh dại được dung hợp vào các vùng xuyên màng và vùng chất nguyên sinh của H5 A/Indonesia/5/2005 được tách dòng thành hệ thống biểu hiện 2X35S-CPMV-HT-PDISP-NOS trong plasmid chứa catxet biểu hiện Plastoxyanin-P19-Plastoxyanin dưới đây sử dụng phương pháp gắn trên cơ sở PCR được thể hiện bởi Darveau và cộng sự (Methods in Neuroscience 26: 77-85 (1995)). Trong chu kỳ PCR thứ nhất, đoạn chứa vùng ngoại bào Rab G mà không cần peptit tín hiệu tự nhiên, vùng xuyên màng và vùng tế bào chất được phóng đại bằng cách sử dụng môi IF-RabG-S2+4.c (Fig.8B, SEQ ID NO:32) và RabG+H5dTm.r (Fig.8C, SEQ ID NO:33), sử dụng gen Rab G được tổng hợp (tương ứng với nt3317-4891 từ số truy cập ngân hàng gen EF206707) (Fig.8D, SEQ ID NO:34) làm khuôn. Đoạn thứ hai chứa vùng xuyên màng và vùng tế bào chất của H5 A/Indonesia/5/2005 được phóng đại bằng cách sử dụng các môi IF-H5dTm.c (Fig.8E, SEQ ID NO:35) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15), sử dụng sản phẩm thiết kế số 972 (xem Fig.94, SEQ ID NO: 134 trong WO 2010/003225, được kết hợp ở đây để tham khảo, với trình tự của sản phẩm thiết kế số 972) làm khuôn. Sản phẩm của phản ứng PCR từ cả hai lần phóng đại tiếp đó được trộn và sử dụng làm khuôn cho chu kỳ phóng đại thứ hai sử dụng IF-RabG-S2+4.c (Fig.8B, SEQ ID NO:32) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15) làm môi. Đoạn tạo thành được tách dòng trong khung bằng

peptit tín hiệu PDI ở cỏ linh lăng trong hệ thống biểu hiện 2X35S-CPMV HT-NOS sử dụng hệ thống tách dòng In-Fusion bởi Clontech (Mountain View, CA). Sản phẩm thiết kế 141 (Fig.8F, 8G) được tiêu hóa bằng enzym giới hạn SbfI và StuI và được sử dụng cho phản ứng lắp ráp In-Fusion. Sản phẩm thiết kế số 141 là plasmit chất nhận được dùng để tách dòng “In Fusion” các gen đang nói đến trong catxet biểu hiện trên cơ sở CPMV HT. Nó cũng kết hợp sản phẩm thiết kế gen để biểu hiện đồng thời gen ức chế TBSV P19 để bất hoạt dưới trình tự khởi đầu và trình tự kết thúc của gen Plastoxyanin ở cỏ linh lăng. Vector là plasmit trên cơ sở pCAMBIA và trình tự đường biên t-ADN từ trái sang phải được thể hiện trên Fig.8F (SEQ ID NO:36). Sự thể hiện dưới dạng sơ đồ vector 141 được thể hiện trên Fig.8G. Sản phẩm thiết kế tạo thành là số 1074 đã biết (Fig.8H, SEQ ID NO: 37). Trình tự axit amin của PDISP - Rab G - A/Indonesia/5/2005 H5 TM+CT được thể hiện trên Fig.8I (SEQ ID NO:38). Sự thể hiện plasmit 1074 được thể hiện trên Fig.8A. Sản phẩm thiết kế này không mã hóa protein M1.

Vùng xuyên màng và đuôi tế bào chất (TM+CT)-NOS (RabG)+H5 A/Indonesia/5/2005 glycoprotein G bệnh dại D-2X35S-PDISP thành BeYDV + hệ thống phóng đại tái bản (Sản phẩm thiết kế số 1094; Fig.8L, 8M)

Trình tự mã hóa vùng ngoại bào glycoprotein G bệnh dại (Rab G) được dung hợp vào vùng xuyên màng và vùng chất nguyên sinh của H5 A/Indonesia/5/2005 được tách dòng thành 2X35S-CPMV-HT-PDISP-NOS thành BeYDV+hệ thống biểu hiện replicaza trong plasmit chứa catxet biểu hiện Plastoxyanin-P19-Plastoxyanin dưới đây sử dụng phương pháp gắn trên cơ sở PCR được thể hiện bởi Darveau và cộng sự (Methods in Neuroscience 26: 77-85 (1995)). Trong chu kỳ PCR thứ nhất, đoạn chứa vùng ngoại bào Rab G mà không có peptit tín hiệu tự nhiên, vùng xuyên màng và vùng tế bào chất được phóng đại bằng cách sử dụng các môi IF-RabG-S2+4.c (Fig.8B, SEQ ID NO:32) và RabG+H5dTm.r (Fig.8C, SEQ ID NO: 33), sử dụng gen Rab G được tổng hợp (tương ứng với nt 3317-4891 từ số truy cập ngân hàng gen EF206707) (Fig.8D, SEQ ID NO:34) làm khuôn. Đoạn thứ hai chứa vùng xuyên màng và vùng tế bào chất của H5 A/Indonesia/5/2005 được phóng đại bằng cách sử dụng các môi IF-H5dTm.c (Fig.8E, SEQ ID NO: 35) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15), sử dụng sản phẩm thiết kế số 972 (xem Fig.94, SEQ ID NO: 134 trong WO 2010/003225, được kết hợp ở đây để tham khảo, với trình tự của sản phẩm thiết kế số 972) làm khuôn. Sản phẩm của

phản ứng PCR từ cả hai lần phóng đại tiếp đó được trộn và sử dụng làm khuôn cho chu kỳ phóng đại thứ hai sử dụng IF-RabG-S2+4.c (Fig.8B, SEQ ID NO:32) và IF-H5dTM.r (Fig.3C, SEQ ID NO: 15) làm mồi. Đoạn tạo thành được tách dòng trong khung bằng peptit tín hiệu PDI ở cở linh lăng trong catxet biểu hiện 2X35S-CPMV HT-NOS thành BeYDV+hệ thống phóng đại replicaza sử dụng hệ thống tách dòng In-Fusion bởi Clontech (Mountain View, CA). Sản phẩm thiết kế 144 được tiêu hóa bằng enzym giới hạn SbfI và StuI và được sử dụng cho phản ứng lắp ráp In-Fusion. Sản phẩm thiết kế số 144 là plasmit chất nhận được dùng để tách dòng “In Fusion” các gen đang nói đến trong catxet biểu hiện trên cơ sở CPMV HT vào hệ thống phóng đại BeYDV. Nó cũng kết hợp sản phẩm thiết kế gen để biểu hiện đồng thời gen ức chế TBSV P19 để bất hoạt dưới trình tự khởi đầu và trình tự kết thúc của gen Plastoxyanin ở cở linh lăng. Vector là plasmit trên cơ sở pCAMBIA và trình tự từ các đường biên t-ADN từ trái qua phải được thể hiện trên Fig.8J (SEQ ID NO:39). Sự thể hiện dưới dạng sơ đồ vector 144 được thể hiện trên Fig.8K. Sản phẩm thiết kế tạo thành là số 1094 đã biết (Fig.8L, SEQ ID NO:40). Trình tự axit amin của PDISP - Rab G - A/Indonesia/5/2005 H5 TM+CT được thể hiện trên Fig.8I (SEQ ID NO:38). Sự thể hiện plasmit 1094 được thể hiện trên Fig.8M. Sản phẩm thiết kế này không mã hóa protein M1.

Ví dụ 5: Sự biểu hiện protein bệnh đại thể khảm ở thực vật

Protein G được biểu hiện In fusion với PDI Sp (sản phẩm thiết kế 1071). Sản phẩm thiết kế 1074 là sự dung hợp của protein G bệnh đại với PDI Sp và vùng H5A/Indo TM+CT. Sản phẩm thiết kế 1094 là sự dung hợp của protein G bệnh đại với vùng BeYDV + rep, PDI Sp và H5A/ Indo TM+CT. Sản phẩm thiết kế 1091 là sự dung hợp của protein G bệnh đại với PDI Sp và BeYDV + rep.

Cây *Nicotiana benthamiana* được phát triển từ hạt giống trong vùng đất phẳng chứa đầy chất 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.

Agrobacteria được gây nhiễm bằng mỗi sản phẩm thiết kế (sản phẩm thiết kế 1071, 1071, 1074, 1091, và 1094) được phát triển trong môi trường YEB có bổ sung 10mM axit 2-(N-morpholino)etansulfonic (MES), 20µM axetosyringon, 50µg/ml kanamycin và 25µg/ml carbenixilin 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 (10mM $MgCl_2$ và 10mM MES pH=5,6) và được bảo quản qua đêm ở nhiệt độ 4°C. Vào ngày thâm nhiễm, các mẻ canh trường được pha loãng trong 2,5 thể tích canh trường và được để ấm trước khi sử dụng. Toàn bộ cây *N. benthamiana* được đặt lộn ngược xuống trong huyền phù vi khuẩn trong thùng thép không gỉ kín không khí dưới chân không có áp suất 20-40 Torr trong thời gian 2-phút. Cây được đưa trở lại nhà kính trong thời gian từ 2 đến 6 ngày ủ cho đến khi thu hoạch. Trừ khi có quy định khác, tất cả các lần thâm nhiễm được tiến hành là thâm nhiễm đồng thời bằng chủng AGL1/172 với tỷ lệ 1:1.

Sau khi ủ, phần trên không của cây được thu hoạch, được làm đông lạnh ở nhiệt độ -80°C và nghiền thành các mẫu. Protein hòa tan tổng số được chiết bằng cách làm đồng nhất (Polytron) mỗi mẫu vật liệu thực vật được nghiền lạnh trong 3 thể tích 50mM Tris pH=8,0 lạnh, 0,15M NaCl, 0,1% Triton X-100 và 1mM phenylmetansulfonyl florua. Sau khi đồng nhất hóa, huyền phù đặc được ly tâm ở 10000g trong thời gian 10 phút ở nhiệt độ 4°C và dịch chiết thô được lọc này (phần nổi trên mặt) được giữ lại 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. Protein được tách ra bằng SDS-PAGE và được chuyển điện tử lên trên màng polyvinyliden diflorua (polyvinylene difluoride - 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.

Phương pháp thẩm tách miễn dịch được tiến hành bằng cách ủ với 0,5ug/ul kháng thể sơ cấp Santa Cruz SE-57995 trong 2% sữa không kem trong TBS-Tween 20 0,1%. Phát hiện sự phát quang bằng phản ứng hóa học được tiến hành sau khi ủ với kháng thể thứ cấp 115-035-146, JIR, chuột kháng dê được tiếp hợp với peroxidaza được pha loãng 1:10000 trong 2% sữa không kem 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).

Như được thể hiện trên Fig.8N, protein G bệnh đại được biểu hiện In fusion với PDI Sp (sản phẩm thiết kế 1071), BeYDV + rep (sản phẩm thiết kế 1094 hoặc 1091), vùng H5A/ Indo TM+CT (sản phẩm thiết kế 1074) hoặc hỗn hợp của chúng.

Ví dụ 6: Protein G bệnh đại thể khả lắp ráp thành hạt tương tự virus

Khả năng của protein G bệnh đại thể khả lắp ráp thành VLP ở thực vật khi không có mặt protein lõi hoặc protein nền cũng được thử nghiệm. Protein chiết từ thực vật được biến nạp bằng cách sử dụng sản phẩm thiết kế 1074 hoặc sản phẩm thiết kế 1094, được tinh sạch và phụ thuộc vào phương pháp sắc ký lọc gel sau đó bằng cách phân tích các phần rửa giải với hàm lượng G ở bệnh đại sử dụng các kháng thể sơ cấp Santa Cruz SE-57995. Phương pháp thẩm tách Western được thể hiện trên Fig.8O thể hiện rằng dịch chiết từ các cây được thâm nhiễm, hàm lượng G bệnh đại thể khả trong các phần rửa giải đạt cực đại trong các phần từ 8 đến 14 với protein được tạo ra bằng cách sử dụng sản phẩm thiết kế 1074, và các phần chính của protein rửa giải trong các phần từ 8 đến 12 với dịch chiết protein được điều chế bằng cách sử dụng sản phẩm thiết kế 1094, chỉ ra rằng sự lắp ráp chúng trong cấu trúc có trọng lượng phân tử rất lớn lớn hơn 2 triệu dalton. Các kết quả này chỉ ra rằng G bệnh đại lắp ráp thành cấu trúc có trọng lượng phân tử cao có kích thước vượt ra ngoài kích thước có trọng lượng phân tử mong muốn của tam phân đồng nhất.

Fig.9A thể hiện phân tích thẩm tách miễn dịch của protein G bệnh đại được biểu hiện từ sản phẩm thiết kế 1074. Fig.9B thể hiện ảnh hiển vi truyền điện tử của protein G bệnh đại VLP được tinh chế có nguồn gốc từ sự biểu hiện sản phẩm thiết kế 1074 thể hiện hình thái học của VLP.

Do đó, G bệnh đại được tích tụ ở mức cao trong các cây được thâm nhiễm agro và được lắp ráp thành hạt tương tự virus khi không có mặt protein lõi hoặc protein nền. Các kết quả này cũng chứng minh rằng sự dimer hóa của vùng TM/CT của HA bệnh cúm thành kháng nguyên không phải ở bệnh cúm, như protein G bệnh đại, đủ để lắp ráp và giải phóng VLP thể hiện kháng nguyên không phải ở bệnh cúm.

Ví dụ 7: Sự lắp ráp catxet biểu hiện với SARS

Glycoprotein S (SARS gS) của virus do hội chứng hô hấp cấp nghiêm trọng B-2X35S-CPMV HT+ vùng xuyên màng và đuôi tế bào chất (TM+CT) H5 A/Indonesia/5/2005 -NOS (Sản phẩm thiết kế số 916; Fig.12A, 12F)

Trình tự mã hóa vùng ngoại bào SARS glycoprotein S (SARS gS) được dung hợp vào các vùng xuyên màng và vùng chất nguyên sinh của H5 A/Indonesia/5/2005 được tách dòng thành hệ thống biểu hiện 2X35S-CPMV-HT dưới đây sử dụng phương pháp gắn trên cơ sở PCR được thể hiện bởi Darveau và cộng sự (Methods in Neuroscience 26: 77-85 (1995)). Trong chu kỳ PCR thứ nhất, đoạn chứa vùng ngoại bào SARS gS mà không có vùng xuyên màng tự nhiên và vùng tế bào chất được phóng đại bằng cách sử dụng môi IF-wtSp-SARSgS.c (Fig.12B, SEQ ID NO:26) và IF-H5dTm+SARSgS.r (Fig.12C, SEQ ID NO:27), sử dụng gen SARS gS được tổng hợp (tương ứng với nt21492-25259 từ số truy cập ngân hàng gen AY278741.1; Fig.12D, SEQ ID NO:28) làm khuôn. Đoạn thứ hai chứa vùng xuyên màng và vùng tế bào chất của H5 A/Indonesia/5/2005 được phóng đại bằng cách sử dụng môi SARSgS+H5dTm.c (Fig.12E, SEQ ID NO:29) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15), sử dụng sản phẩm thiết kế số 972 (xem Fig.94, SEQ ID NO: 134 trong WO 2010/003225, được kết hợp ở đây để tham khảo, với trình tự của sản phẩm thiết kế số 972) làm khuôn. Sản phẩm của phản ứng PCR từ cả hai lần phóng đại tiếp đó được trộn và được sử dụng làm khuôn cho chu kỳ phóng đại thứ hai sử dụng IF-wtSp-SARSgS.c (Fig.12B, SEQ ID NO:26) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15) làm môi. Sản phẩm của phản ứng PCR thứ hai tiếp đó được tiêu hóa bằng ApaI và StuI và được tách dòng thành sản phẩm thiết kế số 995 (Fig.1F, SEQ ID NO: 6) được tiêu hóa bằng enzym giới hạn ApaI và StuI. Sản phẩm thiết kế tạo thành là số 916 đã biết (Fig.12F, SEQ ID NO:30). Trình tự axit amin của SARS gS - A/Indonesia/5/2005 H5 TM+CT được thể hiện trên Fig.12G (SEQ ID NO:31). Sự thể hiện plasmid 916 được thể hiện trên Fig.12A. Sản phẩm thiết kế này không mã hóa protein M1.

Ví dụ 8: Các thử nghiệm biểu hiện với SARS thể khảm có và không có các yếu tố tăng cường sản xuất ở thực vật

Protein SARS được biểu hiện bằng cách sử dụng sản phẩm thiết kế 916 bao gồm gen SARS gS với peptit tín hiệu loại đại và vùng xuyên màng và vùng chất nguyên sinh H5A/ Indo.

Cây *Nicotiana benthamiana* được phát triển từ hạt giống trong vùng đất phẳng chứa đầy chất 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, nhỏ

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.

Agrobacteria được gây nhiễm với mỗi sản phẩm thiết kế (sản phẩm thiết kế 916) được phát triển trong môi trường YEB được bổ sung 10mM axit 2-(N-morpholino)etansulfonic (MES), 20 μ M axetosyringon, 50 μ g/ml kanamycin và 25 μ g/ml carbenixilin 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 (10mM MgCl₂ và 10mM MES pH=5,6). và được bảo quản qua đêm ở nhiệt độ 4°C. Vào ngày thâm nhiễm, các mẻ canh trường được pha loãng trong 2,5 thể tích canh trường và được để ấm trước khi sử dụng. Toàn bộ cây *N. benthamiana* được đặt lộn ngược xuống trong huyền phù vi khuẩn trong thùng thép không gỉ kín không khí dưới chân không có áp suất 20-40 Torr trong thời gian 2-phút. Cây được đưa trở lại nhà kính trong thời gian từ 2 đến 6 ngày ủ cho đến khi thu hoạch. Trừ khi có quy định khác, tất cả các lần thâm nhiễm được tiến hành là thâm nhiễm đồng thời bằng chủng AGL1/172 với tỷ lệ 1:1.

Sau khi ủ, phần trên không của cây được thu hoạch, được làm đông lạnh ở nhiệt độ -80°C và nghiền thành các mẫu. Protein hòa tan tổng số được chiết bằng cách làm đồng nhất (Polytron) mỗi mẫu vật liệu thực vật được nghiền lạnh trong 3 thể tích 50mM Tris pH=8,0 lạnh, 0,15M NaCl, 0,1% Triton X-100 và 1mM phenylmetansulfonyl florua. Sau khi đồng nhất hóa, huyền phù đặc được ly tâm ở 10000g trong thời gian 10 phút ở nhiệt độ 4°C và dịch chiết thô được lọc này (phần nổi trên mặt) được giữ lại 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. Protein được tách ra bằng SDS-PAGE và được chuyển điện tử lên trên màng polyvinylen diflorua (polyvinylene difluoride - 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.

Phương pháp thẩm tách miễn dịch được tiến hành bằng cách ủ 2 μ g/ul Imgenex. Kháng thể sơ cấp IMG-690 trong 2% sữa nguyên kem trong TBS-Tween 20 0,1%. Phát hiện sự phát quang bằng phản ứng hóa học được tiến hành sau khi ủ với kháng

thể thứ cấp 115-035-144, JIR, IgG thử kháng chuột được tiếp hợp với peroxidaza được pha loãng 1:10000 trong 2% sữa nguyên kem 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).

Fig.12H thể hiện phân tích thẩm tách miễn dịch về sự biểu hiện của protein S ở virus do hội chứng hô hấp cấp nghiêm trọng (severe acute respiratory syndrome - SARS) ở cây thuốc lá. Sự biểu hiện của sản phẩm thiết kế 916 (gen SARS gS với peptit tín hiệu loại đại và vùng xuyên màng và vùng đuôi chất nguyên sinh H5A/ Indo) được nghiên cứu.

Ví dụ 9: Sự lắp ráp catxet biểu hiện với protein VZV

Glycoprotein E virus gây bệnh thủy đậu (VZVgE) A-2X35S-CPMV HT + vùng xuyên màng và vùng tế bào chất H5 A/Indonesia/5/2005 (TM+CT)-NOS (Sản phẩm thiết kế số 946)

Trình tự mã hóa vùng ngoại bào glycoprotein E VZV (VZV gE) được dung hợp thành các vùng xuyên màng và vùng chất nguyên sinh của H5 A/Indonesia/5/2005 được tách dòng thành hệ thống biểu hiện 2X35S-CPMV-HT dưới đây sử dụng phương pháp nổi trên cơ sở PCR được thể hiện bởi Darveau và cộng sự (Methods in Neuroscience 26: 77-85 (1995)). Trong chu kỳ PCR thứ nhất, đoạn chứa vùng ngoại bào VZV gE mà không có vùng xuyên màng tự nhiên và vùng tế bào chất được phóng đại bằng cách sử dụng các môi IF-wtSp-VZVgE.c (Fig.10A, SEQ ID NO:20) và IF-H5dTm+VZVgE.r (Fig.10B, SEQ ID NO:21), sử dụng gen VZV gE được tổng hợp (tương ứng với nt3477-5348 từ số truy cập ngân hàng gen AY013752.1; Fig.10C, SEQ ID NO:22) làm khuôn. Đoạn thứ hai chứa vùng xuyên màng và vùng tế bào chất của H5 A/Indonesia/5/2005 được phóng đại bằng cách sử dụng các môi VZVgE+H5dTm.c (Fig.10D, SEQ ID NO:23) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15), sử dụng sản phẩm thiết kế số 972 (xem Fig.94, SEQ ID NO: 134 trong WO 2010/003225, được kết hợp ở đây để tham khảo, với trình tự của sản phẩm thiết kế số 972) làm khuôn. Sản phẩm của phản ứng PCR từ cả hai lần phóng đại tiếp đó được trộn và được sử dụng làm khuôn cho lần phóng đại thứ hai sử dụng IF-wtSp-VZVgE.c (Fig.10A, SEQ ID NO:20) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15) làm môi. Sản phẩm của phản ứng PCR thứ hai tiếp đó được tiêu hóa bằng ApaI và StuI và được tách dòng thành sản phẩm thiết kế số 995 (Fig.1F, SEQ ID NO: 6) được tiêu hóa bằng các enzym giới hạn

ApaI và StuI. Sản phẩm thiết kế tạo thành là số 946 đã biết (Fig.A5, SEQ ID NO: A5). Trình tự axit amin của VZV gE- A/Indonesia/5/2005 H5 TM+CT được thể hiện trên Fig.10F (SEQ ID NO:25). Sự thể hiện plasmit 946 được thể hiện trên Fig.10G. Sản phẩm thiết kế này không mã hóa protein M1.

Ví dụ 10: Sự biểu hiện của các protein VZV thể khảm ở thực vật

Sự biểu hiện của protein E của virus ở bệnh thủy đậu (Varicella Zoster Virus - VZV) chứng minh được việc sử dụng sản phẩm thiết kế 946, bao gồm gen VZV gE với peptit tín hiệu loại đại và vùng H5A/ Indo TM+CT.

Cây *Nicotiana benthamiana* được phát triển từ hạt giống trong vùng đất phẳng chứa đầy chất 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.

Agrobacteria được lây nhiễm với mỗi sản phẩm thiết kế (sản phẩm thiết kế 946) được phát triển trong môi trường YEB có bổ sung 10mM axit 2-(N-morpholino)etansulfonic (MES), 20μM axetosyringon, 50μg/ml kanamycin và 25μg/ml carbenixilin 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 (10mM MgCl₂ và 10mM MES pH=5,6). và được bảo quản qua đêm ở nhiệt độ 4°C. Vào ngày thâm nhiễm, các mẻ canh trường được pha loãng trong 2,5 thể tích canh trường và được để ấm trước khi sử dụng. Toàn bộ cây *N. benthamiana* được đặt lộn ngược xuống trong huyền phù vi khuẩn trong thùng thép không gỉ kín không khí dưới chân không có áp suất 20-40 Torr trong thời gian 2-phút. Cây được đưa trở lại nhà kính trong thời gian từ 2 đến 6 ngày ủ cho đến khi thu hoạch. Trừ khi có quy định khác, tất cả các lần thâm nhiễm được tiến hành là thâm nhiễm đồng thời bằng chủng AGL1/172 với tỷ lệ 1:1.

Sau khi ủ, phần trên không của cây được thu hoạch, được làm đông lạnh ở nhiệt độ -80°C và nghiền thành các mẫu. Protein hòa tan tổng số được chiết bằng cách làm đồng nhất (Polytron) mỗi mẫu vật liệu thực vật được nghiền lạnh trong 3 thể tích 50mM Tris pH=8,0 lạnh, 0,15M NaCl, 0,1% Triton X-100 và 1mM phenylmetansulfonyl florua. Sau khi đồng nhất hóa, huyền phù đặc được ly tâm ở

10000g trong thời gian 10 phút ở nhiệt độ 4°C và dịch chiết thô được lọc này (phần nổi trên mặt) được giữ lại 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. Protein được tách ra bằng SDS-PAGE và được chuyển điện tử lên trên màng polyvinyliden diflorua (polyvinylene difluoride - 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.

Phương pháp thấm tách miễn dịch được tiến hành bằng cách ủ với 1 µg/ul protein kháng VZVgE mAb ở chuột (abcam, ab52549) làm kháng thể sơ cấp trong 2% sữa không kem trong TBS-Tween 20 0,1%. Sự phát hiện sự phát quang bằng phản ứng hóa học được tiến hành sau khi ủ với 115-035-146, JIR, chuột kháng dê được tiếp hợp với peroxidazaza làm kháng thể thứ cấp được pha loãng 1:10000 trong 2% sữa không kem 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).

Fig.11A thể hiện thử nghiệm thấm tách miễn dịch về sự biểu hiện của protein E của virus gây bệnh thủy đậu (Varicella Zoster Virus - VZV) trong các làn 7 đến 9 từ sản phẩm thiết kế 946 (20, 10 và 2 µg dịch chiết tương ứng). Làn 1 đến 5, đối chứng dương - VZV gE tái tổ hợp, 500, 100, 50, 10 và 5ng, tương ứng; làn 6 đối chứng âm.

Khả năng của protein E ở VZV thể khả lắp ráp được thành VLP ở cây khi không có mặt protein lõi hoặc protein nền cũng được nghiên cứu. Protein chiết từ cây được biến nạp sử dụng sản phẩm thiết kế 946 được lọc và phụ thuộc vào phương pháp sắc ký lọc gel sau đó bằng cách phân tích các phần rửa giải đối với protein E ở VZV sử dụng kháng thể sơ cấp kháng protein VZVgE mAb ở chuột (abcam, ab52549). Phương pháp thấm tách Western được thể hiện trên Fig.11B thể hiện rằng trong các dịch chiết từ các cây được thâm nhiễm, các phần rửa giải protein E ở VZV thể khả đạt cực đại trong các phần từ 10 đến 13, chỉ ra rằng protein E ở VZV lắp ráp trong cấu trúc có trọng lượng phân tử rất cao cao hơn 2 triệu dalton. Các kết quả này chỉ ra rằng protein E ở VZV lắp ráp thành cấu trúc có trọng lượng phân tử cao có kích thước vượt ra ngoài trọng lượng phân tử mong đợi của tam phân đồng nhất.

Do đó, protein E ở VZV được tích tụ ở mức cao trong các cây được thâm nhiễm agro và được lắp ráp thành các hạt tương tự virus khi không có mặt protein lõi hoặc protein nền. Các kết quả này cũng chứng minh rằng sự dung hợp của vùng TM/CT của HA cúm thành kháng nguyên không phải ở bệnh cúm, như protein E VZV, đủ để lắp ráp và giải phóng các VLP trình diện kháng nguyên không phải ở bệnh cúm.

Ví dụ 11. Sự lắp ráp catxet biểu hiện với glycoprotein của virus Ebola thể khảm (glycoprotein - GP)

Vùng xuyên màng và đuôi tế bào chất A-2X35S-CPMV HT-PDISP-Zaire Ebolavirus GP (EboGP)+H5 A/Indonesia/5/2005 (TM+CT)-NOS (sản phẩm thiết kế số 1366)

Trình tự mã hóa vùng ngoại bào của glycoprotein của virus Ebola (glycoprotein - GP) từ chủng Zaire 1995 được dung hợp vào vùng xuyên màng và vùng chất nguyên sinh của H5 A/Indonesia/5/2005 được tách dòng thành hệ thống biểu hiện 2X35S-CPMV HT-PDISP-NOS trong plasmit chứa catxet biểu hiện Plastoxyanin-P19-Plastoxyanin dưới đây sử dụng phương pháp gắn trên cơ sở PCR được thể hiện bởi Darveau và cộng sự (Methods in Neuroscience 26: 77-85 (1995)). Gen Ebola GP được tối ưu hóa để sử dụng codon và có hàm lượng GC từ trình tự gen loại đại (tương ứng với nt 6039-8069 từ số truy cập ngân hàng gen AY354458). Vị trí nổi bí ẩn, trình tự Shine-Delgarno, trình tự khởi đầu định ARN và vị trí đi vào ribosom của sinh vật nhân nguyên thủy tiếp đó loại bỏ từ trình tự được tối ưu hóa để tránh cấu trúc hoặc trình tự không mong muốn. Trong chu kỳ PCR thứ nhất, đoạn gen GP được tối ưu hóa chứa trình tự mã hóa vùng ngoại bào (mà không có peptit tín hiệu và vùng xuyên màng và vùng tế bào chất) được phóng đại bằng cách sử dụng các môi IF-Opt_EboGP.s2+4c (Fig.13A, SEQ ID NO: 43) và H5iTMCT+Opt_EboGP.r (Fig.13B, SEQ ID NO: 44), với gen GP được tổng hợp (Fig.13C, SEQ ID NO:45) làm khuôn. Đoạn thứ hai chứa vùng xuyên màng và vùng tế bào chất của H5 từ Influenza A/Indonesia/5/2005 được phóng đại bằng cách sử dụng các môi Opt_EboGP+H5iTMCT.c (Fig.13D, SEQ ID NO: 46) và IF-H5dTm.r (Fig.3C, SEQ ID NO: 15), sử dụng sản phẩm thiết kế số 972 (xem Fig.94, SEQ ID NO: 134 trong WO 2010/003225, được kết hợp ở đây để tham khảo, với trình tự của sản phẩm thiết kế số 972) làm khuôn. Sản phẩm của phản ứng PCR từ cả hai lần phóng đại tiếp đó được trộn và được sử dụng làm khuôn trong chu kỳ phóng đại thứ hai sử dụng IF-Opt_EboGP.s2+4c (Fig.13A, SEQ ID NO: 43) và IF-

H5dTm.r (Fig.3C, SEQ ID NO: 15) làm mẫu. Sản phẩm PCR tạo thành được tách dòng trong khung với peptit tín hiệu PDI ở cở linh lăng trong hệ thống biểu hiện 2X35S-CPMV HT-NOS sử dụng hệ thống tách dòng In-Fusion bởi Clontech (Mountain View, CA). Sản phẩm thiết kế 1192 (Fig.13E) được tiêu hóa bằng enzyme giới hạn SacII và StuI và được sử dụng trong phản ứng lắp ráp In-Fusion. Sản phẩm thiết kế số 1192 là plasmid chất nhận được dùng để tách dòng “In Fusion” các gen đang nói đến trong khung với peptit tín hiệu PDI ở cở linh lăng trong catxet biểu hiện trên cơ sở CPMV HT. Nó cũng kết hợp sản phẩm thiết kế gen để biểu hiện đồng thời gen ức chế TBSV P19 để bất hoạt dưới trình tự khởi đầu và trình tự kết thúc của gen Plastoxyanin ở cở linh lăng. Vector là plasmid trên cơ sở pCAMBIA và trình tự từ các đường biên t-ADN từ trái sang phải được thể hiện trên Fig.13F (SEQ ID NO: 47). Sản phẩm thiết kế tạo thành là số 1366 đã biết (Fig.13G, SEQ ID NO: 48). Trình tự axit amin của PDISP-GP từ Zaire 95 Ebolavirus-H5 TM+CT từ A/Indonesia/5/2005 được thể hiện trên Fig.13H (SEQ ID NO: 49). Sự thể hiện plasmid 1366 được thể hiện trên Fig.13I.

Ví dụ 12. Sự biểu hiện của virus Ebola thể khả năng GP ở thực vật

Sự biểu hiện của virus Ebola (Ebola virus - EV) glycoprotein (GP) chứng minh được rằng việc sử dụng sản phẩm thiết kế 1366, bao gồm gen EV GP với peptit tín hiệu loại đại và vùng H5A/ Indo TM+CT.

Cây *Nicotiana benthamiana* được phát triển từ hạt giống trong vùng đất phẳng chứa đầy chất 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.

Agrobacteria được lây nhiễm với mỗi sản phẩm thiết kế (sản phẩm thiết kế 946) được phát triển trong môi trường YEB có bổ sung 10mM axit 2-(N-morpholino)etansulfonic (MES), 20μM acetosyringon, 50μg/ml kanamycin và 25μg/ml carbenicillin 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 (10mM MgCl₂ và 10mM MES pH=5,6) và được bảo quản qua đêm ở nhiệt độ 4°C. Vào ngày thâm nhiễm, mẻ canh trường được pha loãng trong 2,5 thể tích canh trường và được để làm ấm trước khi sử dụng. Toàn bộ cây *N. benthamiana* được đặt lộn ngược xuống trong huyền phù vi khuẩn trong thùng

thép không gỉ kín không khí dưới chân không có áp suất 20-40 Torr trong thời gian 2-phút. Cây được đưa trở lại nhà kính trong thời gian từ 2 đến 6 ngày ủ cho đến khi thu hoạch.

Sau khi ủ, phần trên không của cây được thu hoạch, được làm đông lạnh ở nhiệt độ -80°C và nghiền thành các mẫu. Protein hòa tan tổng số được chiết bằng cách làm đồng nhất (Polytron) mỗi mẫu vật liệu thực vật được nghiền lạnh trong 3 thể tích 50mM Tris pH=8,0 lạnh, 0,15M NaCl, 0,1% Triton X-100 và 1mM phenylmetansulfonyl florua. Sau khi đồng nhất hóa, huyền phù đặc được ly tâm ở 10000g trong thời gian 10 phút ở nhiệt độ 4°C và dịch chiết thô được lọc này (phần nổi trên mặt) được giữ lại 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. Protein được tách ra bằng SDS-PAGE và được chuyển điện tử lên trên màng polyvinylen diflorua (polyvinylene difluoride - 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 .

Phương pháp thẩm tách miễn dịch được tiến hành bằng cách ủ với 150ng/ml GP Zaire kháng Ebola ở thỏ được tinh chế bằng ái lực (IBT Bioservices, 0301-012) làm kháng thể sơ cấp trong 2% sữa không kem trong TBS-Tween 20 0,1%. Sự phát hiện sự phát huỳnh quang bằng phương pháp hóa học được tiến hành sau khi ủ với kháng thể thứ cấp kháng thỏ ở dê được tiếp hợp peroxidaza (JIR 11-035-144), được pha loãng 1:7500 trong 2% sữa không kem 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). Fig.13J thể hiện phép phân tích thẩm tách miễn dịch để biểu hiện GP virus Ebola thể khảm trong protein chiết từ thực vật được biến nạp bằng sản phẩm thiết kế số 1366. Kết thu được chỉ ra rằng GP Ebola thể khảm được biểu hiện tạm thời.

Tất cả các trích dẫn được kết hợp ở đây để tham khảo. Sáng chế được mô tả về một hoặc nhiều phương án. Tuy nhiên, nó sẽ rõ ràng hơn với những người có hiểu biết trong lĩnh vực kỹ thuật này về số lượng các biến thể và cải biến có thể được thực hiện mà không tách rời khỏi phạm vi của sáng chế như được định rõ trong các điểm yêu cầu bảo hộ.

YÊU CẦU BẢO HỘ

1. Phương pháp sản xuất hạt tương tự virus (VLP) ở thực vật bao gồm các bước:

a) đưa axit nucleic chứa vùng điều hòa hoạt động ở thực vật và được liên kết theo kiểu hoạt động được với trình tự nucleotit thể khảm mã hóa, theo đây, vùng ngoại bào từ protein bề mặt ba đơn phân của virus được dung hợp với vùng xuyên màng và vùng đuôi tế bào chất (TM/CT) cúm, vào thực vật, hoặc bộ phận của thực vật, trong đó TM/CT là từ hemagglutinin cúm và vùng ngoại bào có nguồn gốc từ virus thuộc họ Retroviridae, Rhabdoviridae, Coronaviridae hoặc Filoviridae, và

b) ủ thực vật hoặc bộ phận của thực vật này trong các điều kiện cho phép biểu hiện axit nucleic, nhờ đó tạo ra VLP, VLP này chứa các protein virus, trong đó các protein virus này bao gồm các protein bề mặt virus thể khảm.

2. Phương pháp theo điểm 1, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus có nguồn gốc từ chi Lentivirus, Lyssavirus, Coronavirus hoặc Ebolavirus.

3. Phương pháp theo điểm 1, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus có nguồn gốc từ virus suy giảm miễn dịch ở người (HIV), virus gây bệnh dại, virus gây hội chứng hô hấp cấp nghiêm trọng (SARS), hoặc virus Ebola.

4. Phương pháp theo điểm 1, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus có nguồn gốc từ protein F, protein S, protein env, protein G, glycoprotein vỏ ngoài E, glycoprotein vỏ ngoài B, glycoprotein vỏ ngoài C, glycoprotein vỏ ngoài I, glycoprotein vỏ ngoài H, glycoprotein GP, hoặc hemagglutinin.

5. Phương pháp theo điểm 1, trong đó trong bước ủ (bước b), axit nucleic được biểu hiện tạm thời ở thực vật.

6. Phương pháp theo điểm 1, trong đó, trong bước ủ (bước b), axit nucleic được biểu hiện ổn định ở thực vật.

7. Phương pháp theo điểm 1, trong đó phương pháp này còn bao gồm bước:

c) thu hoạch thực vật và tinh chế VLP.

8. Phương pháp theo điểm 1, trong đó, trong bước đưa vào (bước a), một hoặc nhiều hơn một axit nucleic bổ sung, được chọn từ nhóm bao gồm trình tự nucleotit mã hóa một hoặc nhiều hơn một protein chaperon, protein kênh proton, chất ức chế proteaza, hoặc tổ hợp của chúng, được đưa vào thực vật.

9. Phương pháp theo điểm 1, trong đó VLP không chứa protein nền hoặc protein lõi của virus.
10. Phương pháp theo điểm 1, trong đó trong bước đưa vào (bước a), vùng xuyên màng và vùng đuôi tế bào chất cúm là vùng xuyên màng và vùng đuôi tế bào chất H5, hoặc vùng xuyên màng và vùng đuôi tế bào chất H1, hoặc vùng xuyên màng và vùng đuôi tế bào chất H3.
11. VLP được tạo ra bằng phương pháp theo điểm 1 hoặc điểm 10.
12. VLP theo điểm 11, trong đó VLP này chứa protein virus thể khả mang các N-glycan đặc hiệu ở thực vật hoặc các N-glycan được biến đổi.
13. VLP theo điểm 11, trong đó VLP này chứa một hoặc nhiều hơn một lipid có nguồn gốc từ thực vật.
14. Chế phẩm chứa liều hữu hiệu của VLP theo điểm 11 để gây ra đáp ứng miễn dịch, và chất mang được dùng.
15. Phương pháp theo điểm 1, trong đó vùng xuyên màng và vùng đuôi tế bào chất cúm thu được từ H5 (A/Indonesia/05/2005) và chứa trình tự được xác định trong SEQ ID NO: 41, hoặc vùng xuyên màng và vùng đuôi tế bào chất cúm thu được từ H3 (A/Brisbane/10/2007) và chứa trình tự được xác định trong SEQ ID NO: 42.
16. Phương pháp theo điểm 1, trong đó thực vật là cỏ linh lăng, cải dầu, *Brassica* spp., ngô, *Nicotiana* spp., cỏ linh lăng, khoai tây, nhân sâm, đậu Hà Lan, yến mạch, lúa gạo, đậu tương, lúa mì, lúa mạch, hướng dương hoặc bông.
17. VLP theo điểm 11, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus có nguồn gốc từ chi Lentivirus, Lyssavirus, Coronavirus hoặc Ebolavirus.
18. VLP theo điểm 11, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus có nguồn gốc từ virus suy giảm miễn dịch ở người (HIV), virus gây bệnh dại, virus gây hội chứng hô hấp cấp nghiêm trọng (SARS), hoặc virus Ebola.
19. VLP theo điểm 11, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus có nguồn gốc từ protein F, protein S, protein env, protein G, glycoprotein vỏ ngoài E, glycoprotein vỏ ngoài B, glycoprotein vỏ ngoài C, glycoprotein vỏ ngoài I, glycoprotein vỏ ngoài H, glycoprotein GP hoặc hemagglutinin.

20. VLP theo điểm 11, trong đó VLP này không chứa protein nền hoặc protein lõi của virus.
21. Hạt tương tự virus (VLP) có nguồn gốc thực vật chứa các protein virus, trong đó các protein virus này bao gồm các protein bề mặt virus thể khảm, các protein bề mặt virus thể khảm này chứa vùng ngoại bào từ protein bề mặt ba đơn phân của virus, được dung hợp với vùng xuyên màng và vùng đuôi tế bào chất (TM/CT) cúm, trong đó TM/CT là từ hemagglutinin cúm và vùng ngoại bào có nguồn gốc từ virus thuộc họ Retroviridae, Rhabdoviridae, Coronaviridae hoặc Filoviridae.
22. VLP có nguồn gốc thực vật theo điểm 21, trong đó các protein bề mặt virus thể khảm chứa các N-glycan đặc hiệu ở thực vật hoặc các N-glycan được biến đổi.
23. VLP có nguồn gốc thực vật theo điểm 21, trong đó VLP này chứa một hoặc nhiều hơn một lipid có nguồn gốc từ thực vật.
24. VLP có nguồn gốc thực vật theo điểm 21, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus có nguồn gốc từ chi Lentivirus, Lyssavirus, Coronavirus hoặc Ebolavirus.
25. VLP có nguồn gốc thực vật theo điểm 21, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus có nguồn gốc từ virus suy giảm miễn dịch ở người (HIV), virus gây bệnh dại, virus gây hội chứng hô hấp cấp nghiêm trọng (SARS) hoặc virus Ebola.
26. VLP có nguồn gốc thực vật theo điểm 21, trong đó vùng ngoại bào từ protein bề mặt ba đơn phân của virus có nguồn gốc từ protein F, protein S, protein env, protein G, glycoprotein vỏ ngoài E, glycoprotein vỏ ngoài B, glycoprotein vỏ ngoài C, glycoprotein vỏ ngoài I, glycoprotein vỏ ngoài H, glycoprotein GP, hoặc hemagglutinin.
27. VLP có nguồn gốc thực vật theo điểm 21, trong đó VLP này không chứa protein nền hoặc protein lõi của virus.
28. Chế phẩm để gây ra đáp ứng miễn dịch chứa VLP có nguồn gốc thực vật theo điểm 21 và chất mang được dung.
29. Tế bào thực vật chứa hạt tương tự virus (VLP) chứa các protein virus, các protein virus này bao gồm các protein bề mặt virus thể khảm, các protein bề mặt virus thể

khảm này chứa vùng ngoại bào từ protein bề mặt ba đơn phân của virus, được dung hợp với vùng xuyên màng và vùng đuôi tế bào chất (TM/CT) cúm, trong đó TM/CT là từ hemagglutinin cúm và vùng ngoại bào có nguồn gốc từ virus thuộc họ Retroviridae, Rhabdoviridae, Coronaviridae hoặc Filoviridae.

30. Tế bào thực vật chứa axit nucleic mã hóa protein bề mặt virus thể khảm, axit nucleic này chứa vùng điều hòa hoạt động ở tế bào thực vật và được liên kết theo cách hoạt động được với trình tự nucleotit thể khảm mã hóa, theo đây, vùng ngoại bào từ protein bề mặt ba đơn phân của virus, được dung hợp với vùng xuyên màng và vùng đuôi tế bào chất (TM/CT) cúm, trong đó TM/CT là từ hemagglutinin cúm và vùng ngoại bào có nguồn gốc từ virus thuộc họ Retroviridae, Rhabdoviridae, Coronaviridae hoặc Filoviridae, trong đó việc biểu hiện axit nucleic này tạo ra hạt tương tự virus (VLP), VLP này chứa các protein virus, trong đó các protein virus này bao gồm các protein bề mặt virus thể khảm.

Fig. 1A, SEQ ID Số: 1: Trình tự axit nucleic liên ứng của HIV ConS ΔCFI (Peptit tín hiệu bám sinh được in đậm, vùng xuyên màng bám sinh và vùng chất nguyên sinh được gạch dưới)

ATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCTGATCCTGGGCATGCTG
ATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAC
 ACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGTGCACAACGTGTGGGCCACCCACGCCT
 GCGTGCCCAACCGACCCCAACCCCAAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGA
 ACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCT
 GACCCCTGTGCGTGACCTGAACTGCACCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAG
 GGCGAGATCAAGAACTGCTCCTTCAACATCACCAACGAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGT
 TCTACCGCCTGGACGTGGTGCCATCGACGACAACAACAACAACTCCTCAACTACCGCCTGATCAACTGCAAC
 ACCTCCGCCATCACCCAGGCCTGCCCAAGGTGTCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGG
 CTTGCGCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCAAG
 TGCACCCACGGCATCAAGCCCGTGGTGTCCACCGAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCA
 TCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGACGCTGAACGAGTCCGTGGAGATCAA
 CTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCGGCCAGGCCTTCTACGCCACCGGC
 GACATCATCGGCGACATCCGCCAGGCCACTGCAACATCTCCGGCACCAAGTGAACAAGACCCTGCAGCAGG
 TGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGA
 GATCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTG
 GATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCTGCCCTGCCGCATCAAGCAGATC
 ATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGGCAAGATCACCTGCAAGTCC
 AACATCACCGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCGCCCCG
 GCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGG
 GCGTGGCCCCACCAAGGCCAAGCTGACCGTGACGGCCCGCCAGCTGCTGTCCGGCATCGTGACGAGCAGT
 CCAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGACGCTGACCGTGTGGGGCATCAAGCAGCTGC
 AGGCCCGCGTGTGGCCGTGGAGCGTACCTGAAGGACCGAGCTGCTGGAGATCTGGGACAACATGACCT
 GGATGGAGTGGGAGCGCGAGATCAACAACCTACACCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACC
 AGCAGGAGAAGAACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACCTGGTTCGACATCA
 CCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCC
GTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCTGTCTTCCAGACCCTGATCCCCAACCCCG
CGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCAGGACCGCGACCGCTCCATCCGCCTGG
TGAACGGCTTCTGGCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCTGCGCGAC
TTCATCCTGATCGCCGCCCGACCGTGGAGCTGCTGGGCCGCAAGGGCTGCGCCGCGGCTGGGAGGCCCTG
AAGTACCTGTGGAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAATCCGCCATCTCCCTGCTGGACACCA
CCGCCATCGCCGTGGCCGAGGGCACCGACCGCTGATCGAGGTGGTGCAGCGCGCCTGCCGCGCCATCCTGA
ACATCCCCCGCCGATCCGCCAGGGCCTGGAGCGCGCCTGCTGTAA

Fig. 1B, SEQ ID Số: 2

IF-Apal-SpPDI.c

TGCCCAAATTTGTGGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTTA

Fig. 1C, SEQ ID Số: 3

SpPDI-HIV gp145.r

CACAGTTTCTCGGCGCGAAGATCTGAGAAGGAACCAACACAAGAAGAG

Fig. 1D, SEQ ID Số: 4

IF-SpPDI-gp145.c

TCTCAGATCTTCGCCGCCGAGAACCTGTGGGTGACCGTGTACTAC

Fig. 1E, SEQ ID Số: 5

WtdTm-gp145.r

CCTTTACAGCAGGGCGCGCTCCAGGCCCTGGCGGATGC

Fig. 1F, SEQ ID Số: 6: Caset biểu hiện số 995, từ PacI (ngược chiều với trình tự khởi động) đến AseI (xuôi chiều ngay với trình tự kết thúc NOS). Vị trí giới hạn SbfI đã loại bỏ được in đậm. PDISP-HIV ConS ΔCFI được gạch dưới.

TTAATTAAGTCGACAAGCTTGCATG**CCGGT**CAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCA
AAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAACCTCCTCGG
ATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATC
ATTGCGATAAAGGAAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCAC
GAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGT
GGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACT
TTTCAACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
GTGGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCT
GCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACG
TCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCACTATCCTTCGCA
AGACCTTCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATAATCTTAATAGGTTTTGATAAAAGC
GAACGTGGGGAAACCCGAACCAACCTTCTTCTAACTCTCTCATCTCTTAAAGCAAACCTCTCTCTTGTCTC
TTTCTTGCCTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAAGTACAACGTTTTCTTCACTGAAGCG
AAATCAAAGATCTCTTTGTGGACACGTAGTGCAGGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTCTGGT
TGGTCTTGGGAAAAGAAAGCTTGTCTGGAGGCTGCTGTTAGCCCCATACATTACTTGTACGATTCTGTGACT
TTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTGTCTGA
TTGGTTCTATAAGAAATCTAGTATTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAAGTTGGAGAAAGATTGT
TAAGCTTCTGTATATTCTGCCAAATTTGTGGGGCCATGGCGAAAAACGTTGCGATTTCGGCTTATTGTTTTCT
TCTTCTGTGTTGGTTCCTTCTCAGATCTTCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCCG
TGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGTGCACAACGT
GTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCAAGGAGATCGTGCTGGAGAAGCTGACCGAGAA
CTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAAGTCCCTG
AAGCCCTGCGTGAAGCTGACCCCTGTGCGTGACCTGAAGTGCACCAACGTGAACGTGACCAACACCACCA
ACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGACAAGAAGC
AGAAGGTGTACGCCCTGTTCTACCGCTGGACGTGGTGGCCATCGACGACAACAACAACAACTCCTCCAACCTAC
CGCCTGATCAACTGCAACACCTCCGCCATCACCAGGCCTGCCCAAGGTGTCTTCGAGCCCATCCCCATCCA
CTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGAACGACAAGAAGTTCAACGGCACCGGCCCTGCAAG
AACGTGTCCACCGTGCAGTGACCCACGGCATCAAGCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCC
TGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGAGCTGAA
CGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCGGCCAG
GCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCTCCGGCACCAAGTGGG
ACAAGACCTGACGAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTC
CTCCGGCGGCGACCTGGAGATCACCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCG
GCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAACAACAACAACCAACGACACCATCACCTGCC
CTGCCGATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGG
CAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACCAACGAG
ACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTACAAGGTG
GTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCTGACCGTGCAGGCCCGGCCAGCTGCTGTCC
GGCATCGTGAGCAGCAGTCCAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGT
GGGCGATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGAG
ATCTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCAACAACCTACCCGACATCATCTACTCCCTGA
TCGAGGAGTCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCCTCCCTGT
GGAAGTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGG

CCTGCGCATCGTGTTGCGCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCTGTCTTCCAGA
CCCTGATCCCCAACCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCAGGACCGC
GACCGCTCCATCCGCCTGGTGAACGGCTTCCTGGCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTC
CTACCACCGCCTGCGCGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGGCCGCAAGGGCCTGCGC
CGCGGCTGGGAGGCCCTGAAGTACCTGTGGAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAAGTCCGCC
ATCTCCCTGCTGGACACCACCGCCATCGCCGTGGCCGAGGGCACCGACCGCGTGATCGAGGTGGTGCAGCGC
GCCTGCCGCGCCATCCTGAACATCCCCGCCGATCCGCCAGGGCCTGGAGCGCGCCCTGCTGTAAAGGCCTA
 TTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCAATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGT
 GTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTGTCCTTCAGCAAGGACACAAAAAGA
 TTTTAATTTTATTAATAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTC
 AAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTT
 GAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAG
 TCCCGCAATTATACATTTAATACGCGATAGAAAACAAATATAGCGCGCAAAGTAAATTATCGCGCGC
 GGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCC

Fig. 1G, SEQ ID Số: 7: trình tự axit amin của PDISP-HIV ConS ΔCFI

MAKNVAIFGLLFSLLVLVPSQIFAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNWVATHACVPTDPNP
 QEIVLENTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVNVTNTTNTEEEKGEIKNCSFNI
 TTEIRDKKQKVYALFYRLDVVPIDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTG
 PCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYAT
 GDIIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGN
 GTKNNNNNTNDTITLPCRIKQIINMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMR
 DNWRSELYKYKVVKIEPLGVAPTKAKLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYL
 KDQQLLEIWDNMTWMEWEREINNYTDIISLIEESQNQQEKNEQELLALDKWASLWNWFDITNWLWYIKIFIMIV
 GGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDRPEGIEEGGEQDRDRSIRLVNGFLALAWDDLRLSLCLFSY
 HRLRDFILIAARTVELLGRKGLRRGWEALKYLWNLLQYWGQELKNSAISLLDTTAIAVAEGTDRVIEVVQRACRAILN
 IPRRIRQGLERALL

Fig. 1H: Sự thể hiện dưới dạng sơ đồ của sản phẩm thiết kế số 995

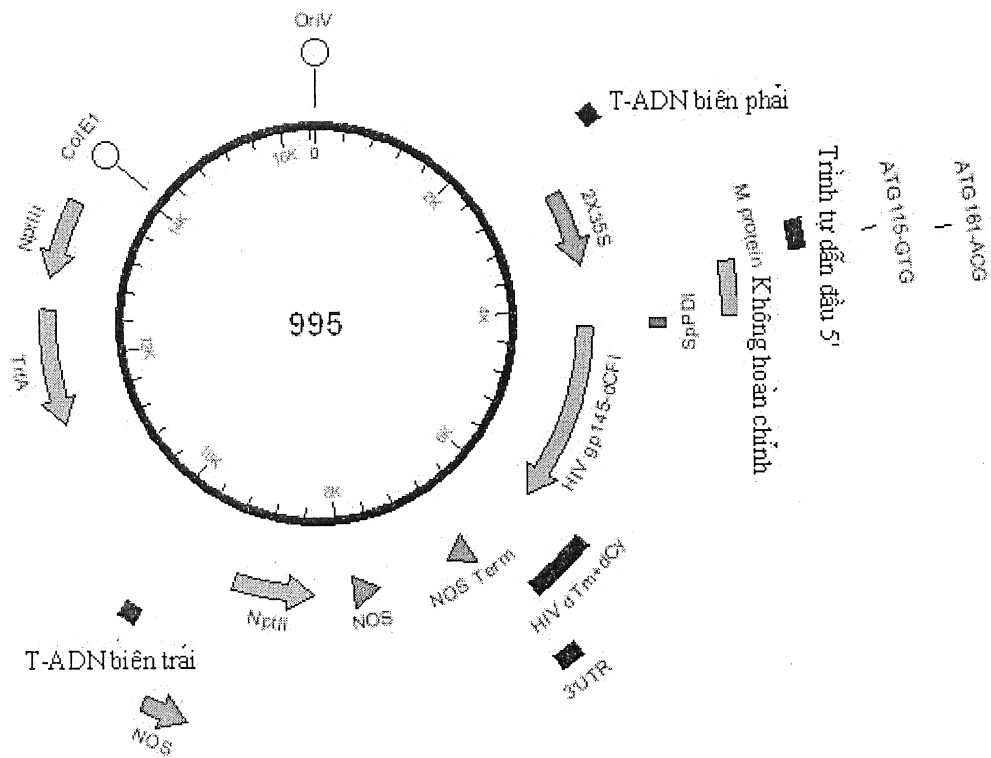


Fig. 2A, SEQ ID Số: 8

IF-H3dTm+gp145.r

CCATAGTATCCAATCGATGTACCACAGCCAGTTGGTGATGTGGAAC

Fig. 2B, SEQ ID Số: 9:

Gp145+H3dTm.c

TGGCTGTGGTACATCGATTGGATACTATGGATTTCCTTTGCATATCAT

Fig. 2C, SEQ ID Số: 10:

H3dTm.r

CCTTCAAATGCAAATGTTGCACCTAATGTTGCCTT

Fig. 2D, SEQ ID Số: 11: Caset biểu hiện số 997, từ PacI (ngược chiều với trình tự khởi động) đến AscI (xuôi chiều ngay với trình tự kết thúc NOS). Vị trí giới hạn SbfI đã loại bỏ được in đậm. PDISP-HIV ConS ΔCFI/Brisbane/10/2007 H3 TM+CT được gạch dưới.

TTAATTAAGTCGACAAGCTTGCATG**CCGGT**CAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCA
AAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGG
ATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATC
ATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCAACCCAC
GAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGT
GGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACT
TTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
GTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCT
GCCGACAGTGGTCCCAAAGATGGACCCCAACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACG
TCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCA
AGACCTTCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATAGTCTTAATAGTCTTCTGATAAAGC
GAACGTGGGGAAACCCGAACCAAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTCTGTC
TTTCTTGCCTGAGCGATCTTCAACGTTGTCAGATCGTCTCGGCACCAAGTACAACGTTTTCTTCTCACTGAAGCG
AAATCAAAGATCTCTTTGTGGACACGTAGTGCAGCGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTG
TGGTCTTGGGAAAAGAAAGCTTGTCTGGAGGCTGCTGTTAGCCCATACATTACTTGTACGATTCTGCTGACT
TTCGCGGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCTGTACTTCTTCTTCTTCTTCTGCTGA
TTGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAAGTGGAGAAAGATTGT
TAAGCTTCTGTATATTCTGCCAAATTTGTGCGGGCCATG**GGCGAAAAACGTTGCGATTTCGGCTTATTGTTTT**
TCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCG
TGTGGAAGGAGGCCAACACCACCTGTTCTGCGCTCCGACGCCAAGGCCTACGACACCGAGGTGCACAACGT
GTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCCAAGGAGATCGTGCTGGAGAACGTGACCGAGAA
CTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCACTCCCTG
AAGCCCTGCGTGAAGCTGACCCCTGTGCGTGACCTGAACTGCACCAACGTGAACGTGACCAACACCACCA
ACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCTTCAACATCACCACCGAGATCCGCGACAAGAAGC
AGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCATCGACGACAACAACAACACTCTCCAACATC
CGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCAAGGTGTCTTCGAGCCCATCCCCATCCA
CTACTGCGCCCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCTGCAAG
AACGTGTCCACCGTGACGTGCACCCACGGCATCAAGCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCC
TGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGACGCTGAA
CGAGTCCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCGGCCAG
GCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCTCCGGCACCAAGTGGA
ACAAGACCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTC
CTCCGCGGCGACCTGGAGATCACCAACCTCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCG
GCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAACAACAACAACCAACGACACCATCACCTGCC
CTGCCGATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGG
CAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGCAACAACAACCAACGAG
ACCGAGATCTTCGCCCCCGCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTACAAGGTG
GTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCTGACCGTGACGGCCCGCCAGCTGCTGTCC
GGCATCGTGACGAGCAGTCCAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGT
GGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGAG
ATCTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCAACAACCTACACCGACATCATCTACTCCCTGA
TCGAGGAGTCCAGAACAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGT
GGAAGTGGTTCGACATCACCAACTGGCTGTGGTACATCGATTGGATACTATGGATTTCTTTGCCATATCATGT
TTTTTGCTTTGTGTTGCTTTGTTGGGGTTCATCATGTGGGCCCTGCCAAAAAGGCAACATTAGGTGCAACATTTG
CATTTGAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTT
TGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTGCTCCCTTCAGCAA
GGACACAAAAAGATTTTAAATTTATTAATAAAAAAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGA
CCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCATGATTATC

ATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGAAGTTATTTATGAGATGGGTT
 TTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAAAAAATATAGCGCGCAAAGTAGGATAA
 ATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCC

Fig. 2E, SEQ ID Số: 12: Trình tự axit amin của PDISP-HIV ConS ΔCFI/Brisbane/10/2007 H3 TM+CT

MAKNVAIFGLFLSLLVLVPSQIFAAENLWVTVYGVVWKEANTTLFCASDAKAYDTEVHNWVATHACVPTDPNP
 QEIVLENVTENFNMWKNMVEQMHEHIIISLWDQSLKPCVKLTPLCVTLNCTNVNVTNTTNNTEEKGEIKNCSENI
 TTEIRDKKQKVYALFYRLDVPIDDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDDKFNKGTG
 PCKNVSTVQCTHGKPVVSTQLLLNGLSAAEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSRIIGPGQAFYAT
 GDIIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLITTHSFNCRGEFFYNTSGLFNSTWIGN
 GTKNNNNTNDTITLPCRIKQIINMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMR
 DNWRSELYKYKVVKIEPLGVAPTAKLTVQARQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYL
 KDQQLLEIWDNMTWMEWEREINNNTDIIYSLIEESQNNQKEKNEQELLALDKWASLWNNWFDITNWLWYIDWILW
 ISFAISCFLLCVALLGFIMWACQKGNIRCNICI

Fig. 2F: Sự thể hiện dưới dạng sơ đồ của sản phẩm thiết kế 997

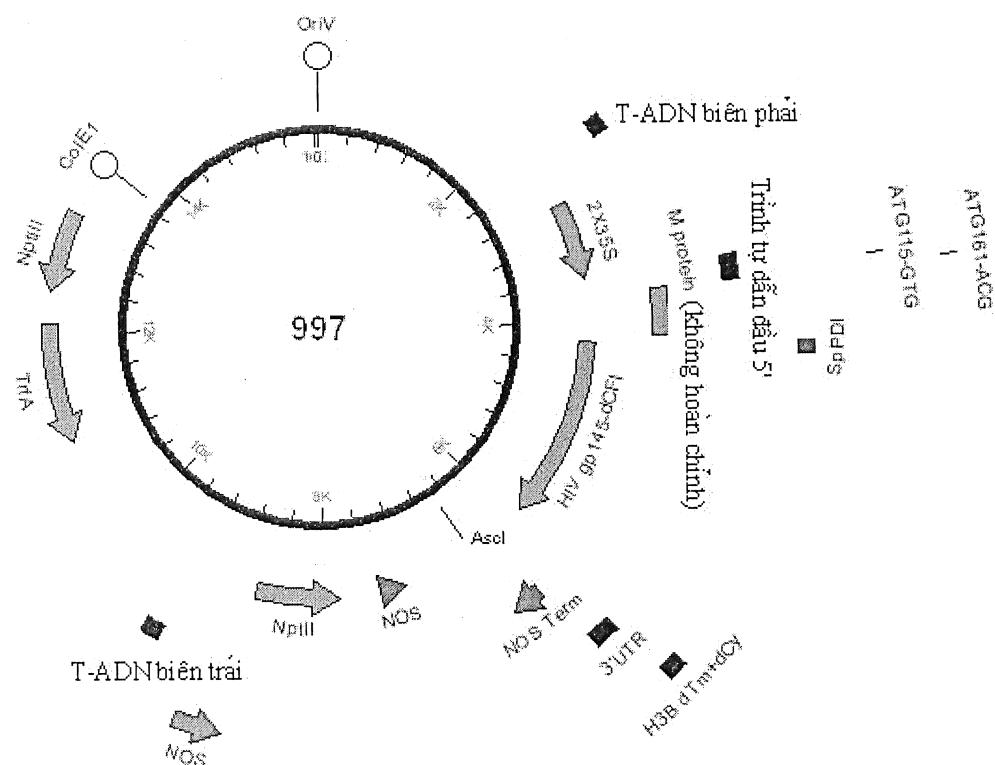


Fig. 3A, SEQ ID Số: 13:

IF-H5dTm+gp145.r

AATTGACAGTATTTGGATGTACCACAGCCAGTTGGTGATGTCTGAAC

Fig. 3B, SEQ ID Số: 14:

Gp145+H5dTm.c

CTGGCTGTGGTACATCCAAATACTGTCAATTTATTCAACAGTGGC

Fig. 3C, SEQ ID Số: 15:

IF-H5dTm.r

CCTTTAAATGCAAATTCTGCATTGTAACGATCCAT

Fig. 3D, SEQ ID Số: 16: Caset biểu hiện số 999, từ PacI (ngược chiều với trình tự khởi động) đến AscI (xuôi chiều ngay với trình tự kết thúc NOS). Vị trí giới hạn SbfI đã loại bỏ được in đậm. PDISP-HIV ConS ΔCFI-A/Indonesia/5/2005 H5 TM+CT được gạch dưới.

TTAATTAAGTCGACAAGCTTGCATG**CCGGT**CAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCA
AAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAACCTCCTCGG
ATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATC
ATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCAC
GAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGT
GGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACT
TTTCAACAAAGGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
GTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCT
GCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACG
TCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCACTATCCTTCGCA
AGACCTTCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTTAAATCTTAATAGGTTTTGATAAAGC
GAACGTGGGGAAACCCGAACCAACCTTCTTCAAACCTCTCTCATCTCTTAAAGCAAACCTCTCTCTTGTCT
TTTCTTGCCTGAGCGATCTTCAACGTTGTGATCGTGCTTCGGCACCAGTACAACGTTTTTCTTCACTGAAGCG
AAATCAAAGATCTTTGTGGACACGTAGTGGCGGCCATTAAATAACGTGTACTTGTCTATTCTTGTCTGGTG
TGGTCTTGGGAAAAGAAAGCTTGTCTGGAGGCTGCTGTTAGCCCCATACATTACTTGTACGATTCTGTCTGACT
TTCGGCGGGTGCAATATCTTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTCTGCTGA
TTGGTCTATAAGAAATCTAGTATTTTCTTGAACAGAGTTTTCCCGTGGTTTTCGAAGTTGGAGAAAGATTGT
TAAGCTTCTGTATATTCTGCCAAATTTGTCGGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTT
TCTTCTTGTGTTGGTTCCTTCTCAGATCTTCGCCGCCGAGAACCTGTGGGTGACCGTGACTACGGCGTGCCCG
TGTGGAAGGAGGCCAACACCACCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGTGCACAACGT
GTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCCCAAGGAGATCGTGCTGGAGAACGTGACCGAGAA
CTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTG
AAGCCCTGCGTGAAGCTGACCCCTGTGCGTGACCTGAAGTGCACCAACGTGAACGTGACCAACACCACCA
ACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTCTTCAACATCACCACCGAGATCCGCGACAAGAAGC
AGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACAACAACAACACTCTCCAACCTAC
CGCCTGATCAACTGCAACACCTCCGCCATCACCAGGCCTGCCCAAGGTGTCCTTCGAGCCCATCCCCATCCA
CTACTGCGCCCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCTGCAAG
AACGTGTCCACCGTGAGTGACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCC
TGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCATCATCGTGCAAGTGA
CGAGTCCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCGCGATCGGCCCGGCCAG
GCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCTCCGGCACCAAGTGGA
ACAAGACCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACCATCATCTTCAAGCCCTC
CTCCGCGGCGACCTGGAGATCACCACCACTCTTCAACTGCCGCGCGAGTCTTCTACTGCAACACCTCCG

GCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCTGCC
CTGCCGCATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGG
CAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAACGAG
ACCGAGATCTTCGCCCCGGCGGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTG
GTGAAGATCGAGCCCCTGGGCGTGGCCCCACCAAGGCCAAGCTGACCGTGCAGGCCCGCCAGCTGCTGTCC
GGCATCGTGCAGCAGCAGTCCAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGTGT
GGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGAG
ATCTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCAACAACACACCGACATCATCTACTCCCTGA
TCGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGT
GGAAGTGGTTCGACATCACCAACTGGCTGTGGTACATCCAAATACTGTCAATTTATTCAACAGTGGCGAGTTCC
CTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAGAATTTG
CATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTGAGCGTTTTT
TGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTGCTCCCTTCAGCAA
GGACACAAAAAGATTTTAATTTATTAATAAAAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGA
CCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATC
ATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTT
TTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAAACAAATATAGCGCGCAAACTAGGATAA
ATTATCGCGCGCGGTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCC

Fig. 3E, SEQ ID Số: 17: Trình tự axit amin của PDISP-HIV ConS ΔCFI-A/Indonesia/5/2005 H5 TM+CT

MAKNVAFGLLFSLLVLVPSQIFAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNWVATHACVPTDPNP
 QEIVLENVTFENFMWKNMVEQMHEIISLWDQSLKPCVKLTPLCVTLNCTNVNVTNTTNNTEKEIKNCSFNI
 TTEIRDKKQKVYALFYRLDVVPIDNNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTG
 PCKNVSTVQCTHGIKPVVSTQLLNGSLAEIIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYAT
 GDIIGDIRQAHCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGN
 GTKNNNNNTNDITLPCRIKQIINMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGDMR
 DNWRSELYKYKVVKIEPLGVAPTAKLTVQARQLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYL
 KDQQLLEIWDNMTWMEWEREINNYTDIIYSLEESQNQQEKNEQELLALDKWASLWNWFDITNWLWYIQILSIYS
 TVASSLALAIMMAGLSLWMCSNGSLQCRICI

Fig. 3F: Sự thể hiện dưới dạng sơ đồ của sản phẩm thiết kế số 999

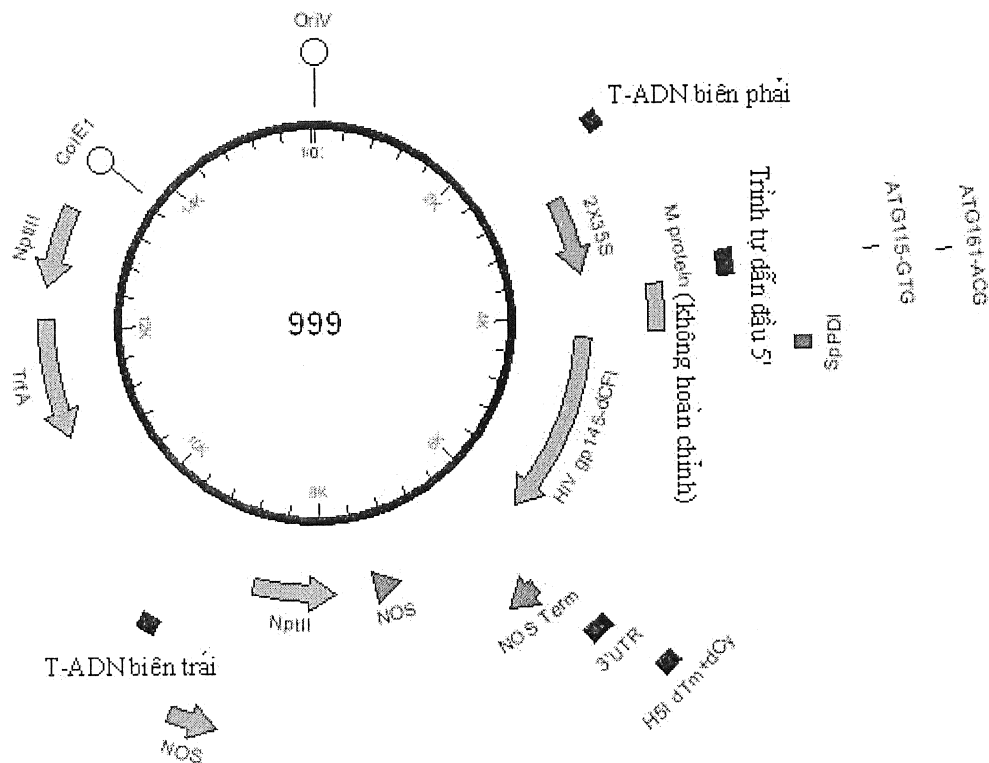


Fig. 4A, SEQ ID Số: 18: Caset biểu hiện số 172, từ XmaI (ngược chiều với trình tự khởi động plastocyanin) đến EcoRI (xuôi chiều ngay với trình tự kết thúc plastocyanin). Trình tự axit nucleic TBSV P19 được gạch dưới.

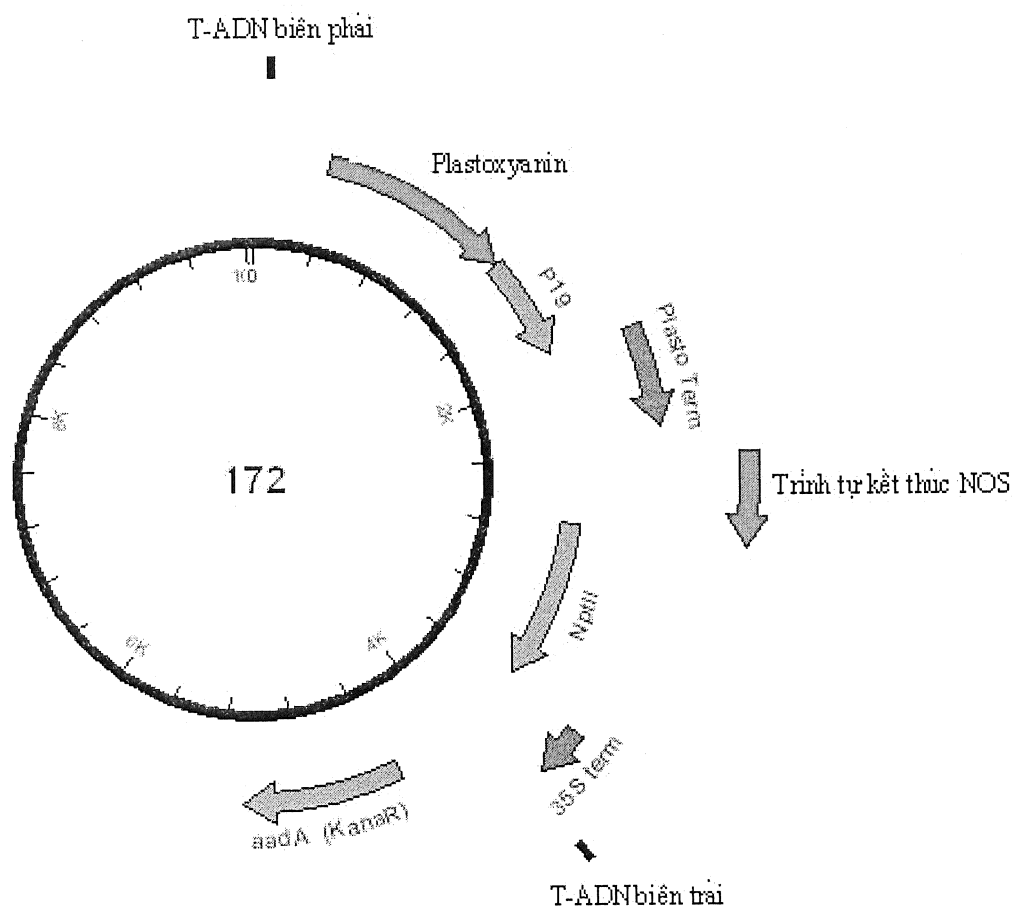
CCCGGGCTGGTATATTATATGTTGTCAAATAACTCAAAAACCATAAAGTTTAAGTTAGCAAGTGTGTACATT
 TTTACTTGAACAAAAATATTACCTACTACTGTTATAAATCATTATTAAACATTAGAGTAAAGAAATATGGATGA
 TAAGAAACAAGAGTAGTGATATTTTGACAACAATTTTGTGCAACATTTGAGAAAATTTGTTGTTCTCTCTTTTC
 ATTTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGAGGGAGAATAAAAAATAATGTGAGTATGAGAGAGAA
 AGTTGTACAAAAGTTGTACCAAAAATAGTTGTACAAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAG
 GGTTAATTGCTGTAATAAATAAGGATGACGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATT
 AAAAAAGAAAGAAATAAATTATTTTAAATTAAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATT
 GATGAAAAGAGTTGGATTAAAGTTGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTC
 CTATATATTGCCCATAGAGTCAGTTAACTCATTTTATATTTTATAGATCAAATAAGAGAAATAACGGTATATT
 AATCCCTCCAAAAAAGGATATTTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCG
 TAGGAGGATAACATCCAATCCAACCAATCAACAATCCTGATGAGATAACCACTTTAAGCCCAAGCATCTGT
 GGCACATCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATC
 ACCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACT
 AATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAAAAGACGCTAGGGAAACAAGCTAA
CAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACCTTCTCCCTTCAAACCTTCTGACGAAAAGTCCGAGTTGG
ACTGAGTGGCGGCTACATAACGATGAGAOGAATTCGAATCAAGATAATCCOCTTGTTTCAAGGAAAAGCTGGG
GTTTCGGGAAAAGTTGTATTAAAGAGATATCTCAGATAOGACAGGACGGAAGCTTCAGTGCACAGAGTCCCTGG
ATCTTGGACGGGAGATTCGGTTAACTATGCAGCATCTCGATTTTTCGTTTCGACCAGATCGGATGTACCTATA
GTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACTCTTCAGCATCTCTGTGAGATGGCA

ATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGGAAAGTAATGTATCAAGAGGATGCCCTG
 AAGSTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTCGAGGGCGAGCTCTAAGTTAAAATGCTTCCTCGTCT
 CCTATTTATAATATGGTTTGTTATTGTTAATTTTGTTCCTGTAGAAGAGCTTAATTAATCGTTGTTGTTATGAAAT
 ACTATTTGTATGAGATGAAGTGGTGTAAATGTAATTCATTTACATAAGTGGAGTCAGAATCAGAATGTTTCCTCC
 ATAATACTAGACATGAAGACCTGCCGCGTACAATTGTCTTATATTTGAACAACATAAAATTGAACATCTTTTGC
 CACAACCTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTTCAATAGATTAATAATGGAAATATCAGTT
 ATCGAAATTCATTAACAATCAACTTAACGTTATTAATACTACTAATTTTATATCATCCCCCTTGATAAATGATAGTAC
 ACCAATTAGGAAGGAGCATGCTCGAGGCCTGGCTGGCCGAATTC

Fig. 4B, SEQ ID Số: 19: Trình tự axit amin của gen ức chế TBSV P19 bị bất hoạt

MERAIQGNDAREQANSERWDGGSGGTTSPFKLPDESPSWTEWRLHNDETNSNQDNPLGFKESWGFGKVFVKRY
 LRYDRTEASLHRVLGSWTGDSDVNYAASRFFGFDQIGCTYSIRFRGVSITVSGGSRTLQHLCEMAIRSKQELLQLAPIEV
 ESNVSRGCPGEGTQTFEKESE

Fig. 4C: Sự thể hiện sản phẩm thiết kế số 172



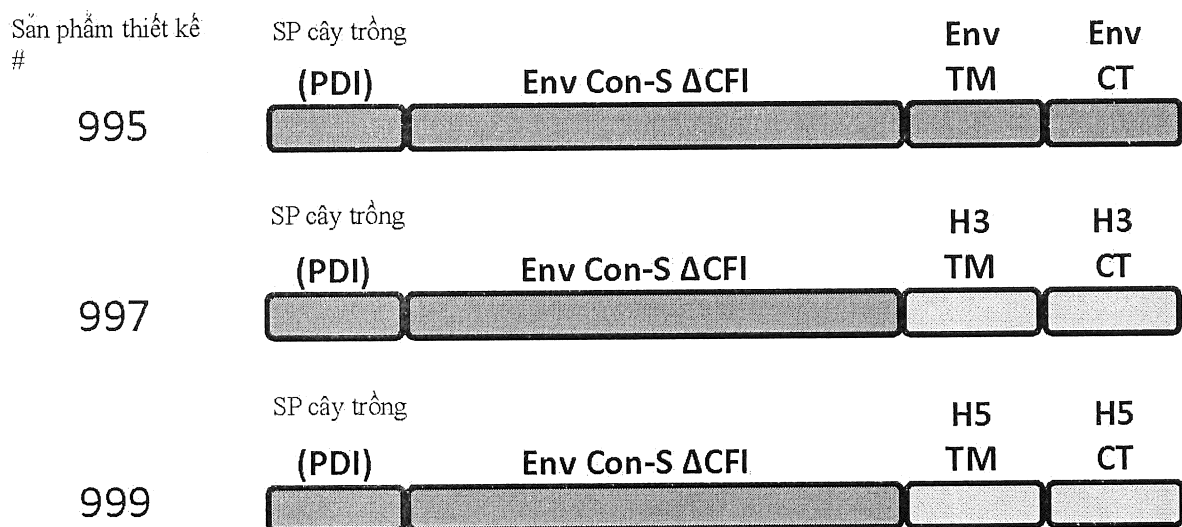


Fig. 5

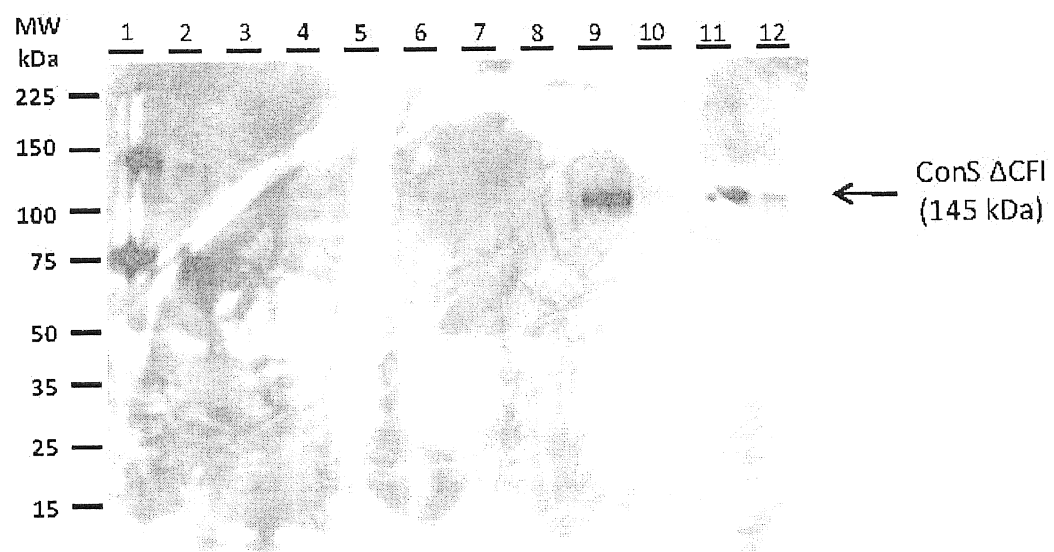


Fig. 6

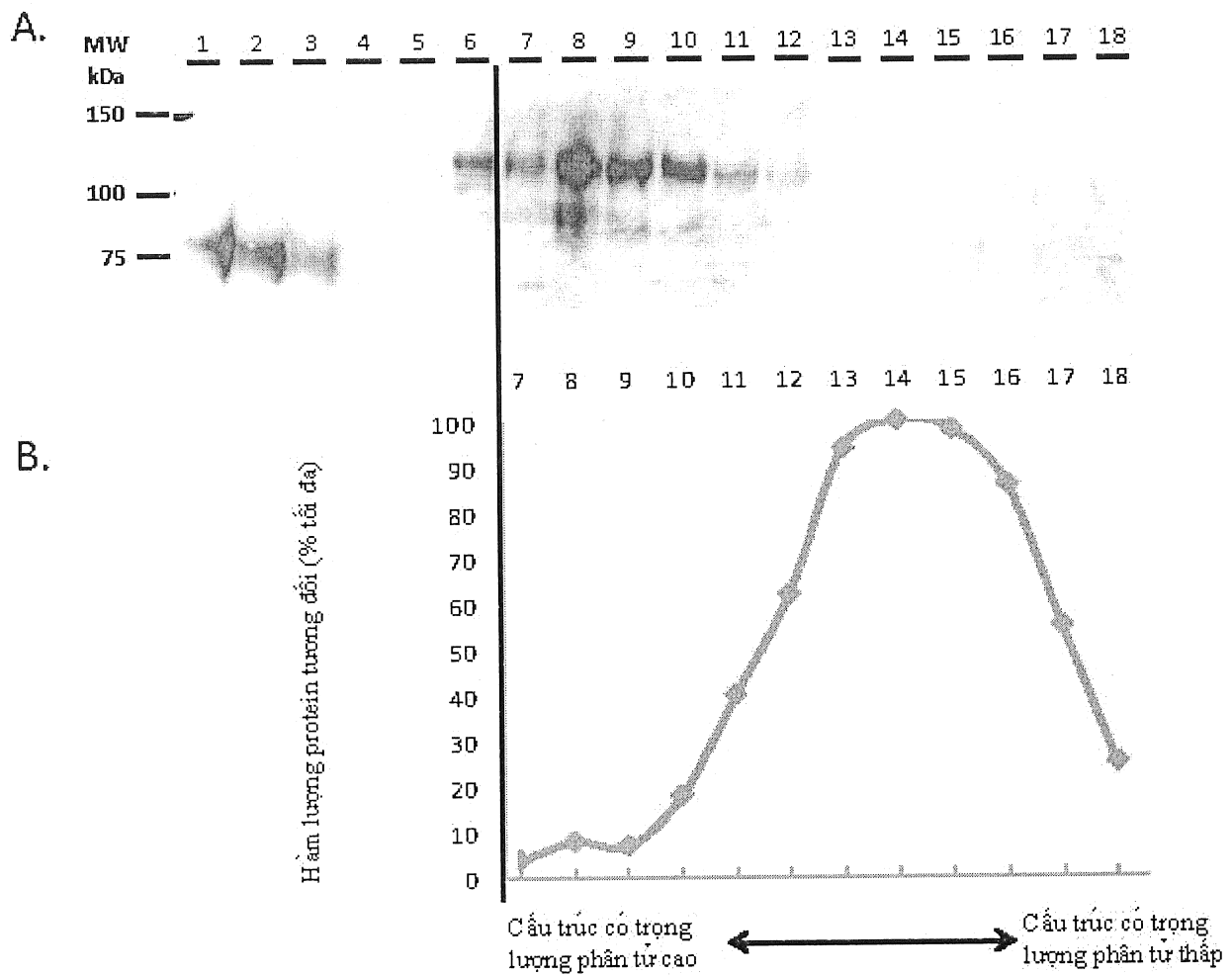


Fig. 7

Fig 8A

Thể hiện dưới dạng sơ đồ sản phẩm thiết kế số 1074

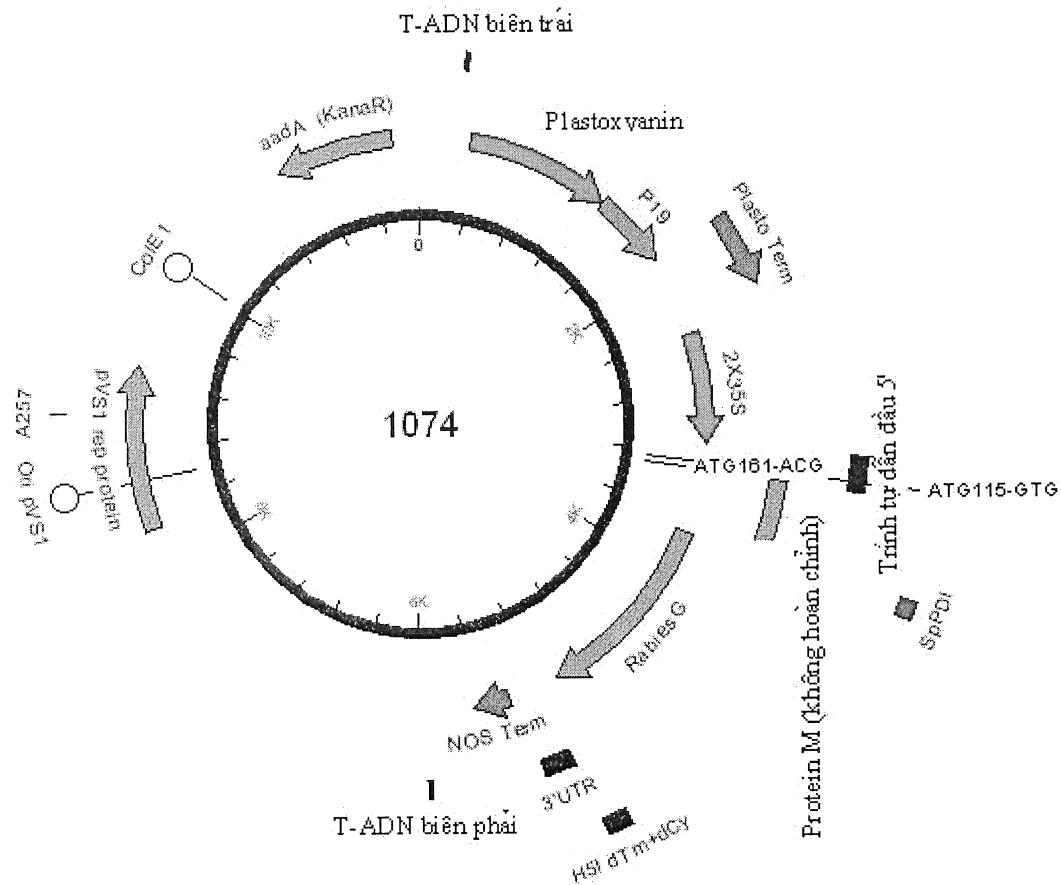


Fig. 8B, SEQ ID Số:32

IF-RabG-S2+4.c

TCTCAGATCTTCGCCAAATTCCTATTTACACGATACCAGAC

Fig. 8C, SEQ ID Số:33

RabG+H5dTm.r

CTGTTGAATAAATTGACAGTATTTGATACTTCCCCAGTTCGGGAGAC

Fig. 8D, SEQ ID Số:34

Gen Rab G được tổng hợp (tương ứng với nt 3317-4891 từ Số gia nhập ngân hàng gen EF206707) (Peptit tín hiệu bầm sinh được in đậm, vùng xuyên màng bầm sinh và vùng chất nguyên sinh được gạch dưới)

ATGGTTCCTCAGGCTCTCCTGTTTGTACCCCTTCTGGTTTTTCCATTGTGTTTTGGGAAATTC
CCTATTTACACGATACCAGACAAGCTTGGTCCCTGGAGCCCGATTGACATACATCACCTCAG
CTGCCCCAAACAATTTGGTAGTGAGGACGAAGGATGCACCAACCTGTCAGGGTTCTCCTAC
ATGGAACCTTAAAGTTGGATACATCTTAGCCATAAAAAATGAACGGGTTCACCTGCACAGGCGT
TGTGACGGAGGCTGAAACCTACACTAACTTCGTTGGTTATGTCACAACCACGTTCAAAAGAA
AGCATTTCCGCCCAACACCAGATGCATGTAGAGCCGCGTACAACCTGGAAGATGGCCGGTGA
CCCCAGATATGAAGAGTCTCTACACAATCCGTACCCTGACTACCGCTGGCTTCGAACTGTAA
AAACCACCAAGGAGTCTCTCGTTATCATATCTCCAAGTGTGGCAGATTTGGACCCATATGAC
AGATCCCTTCACTCGAGGGTCTTCCCTAGCGGGAAGTGCTCAGGAGTAGCGGTGTCTTCTA
CCTACTGCTCCACTAACCAAGATTACACCATTGGATGCCCCGAGAATCCGAGACTAGGGATG
TCTTGTGACATTTTTACCAATAGTAGAGGGAAGAGAGCATCCAAAGGGAGTGAGACTTGCGG
CTTTGTAGATGAAAGAGGCCTATATAAGTCTTTAAAAGGAGCATGCAAACATCAAATGAAACCA
AGTTCTAGGACTTAGACTTATGGATGGAACATGGGTGCGCATGCAAACATCAAATGAAACCA
AATGGTGCCCTCCCGATCAGTTGGTGAACCTGCACGACTTTCGCTCAGACGAAATTGAGCA
CCTTGTGTAGAGGAGTTGGTCAGGAAGAGAGAGGAGTGTCTGGATGCACTAGAGTCCATC
ATGACAACCAAGTCAGTGAGTTTCAGACGTCTCAGTCATTTAAGAAAACCTTGTCCTGGGTTT
GGAAAAGCATATACCATATTCAACAAGACCTTGATGGAAGCCGATGCTCACTACAAGTCAGT
CAGAACTTGGAATGAGATCCTCCCTTCAAAGGGTGTTTAAGAGTTGGGGGGAGGTGTCAT
CCTCATGTGAACGGGGTGTTTTTCAATGGTATAATATTAGGACCTGACGGCAATGTCTTAATC
CCAGAGATGCAATCATCCCTCCTCCAGCAACATATGGAGTTGTTGGAATCCTCGGTTATCCC
CCTTGTGCACCCCTGGCAGACCCGTCTACCGTTTTCAAGGACGGTGACGAGGCTGAGGAT
TTTGTGTAAGTTCACCTTCCCGATGTGCACAATCAGGTCTCAGGAGTTGACTTGGGTCTCCC
GAAGTGGGGGAAGTAT**GTATTACTGAGTGCAGGGGCCCTGACTGCCTTGATGTTGATAATTT**
TCCTGATGACATGTTGTAGAAGAGTCAATCGATCAGAACCTACGCAACACAATCTCAGAGGG
ACAGGGAGGGAGGTGTCAGTCACTCCCCAAAGCGGGAAGATCATATCTTCATGGGAATCAC
ACAAGAGTGGGGGTGAGACCAGACTGTGA

Fig. 8E, SEQ ID Số:35

IF-H5dTm.c

CAAATACTGTCAATTTATTCAACAGTGGC

Fig. 8F, SEQ ID Số:36

Sản phẩm thiết kế 141 từ t-ADN từ trái sang phải. Hệ thống biểu hiện 2X35S-CPMV-HT-PDISP-NOS với caset biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTTTTA
 ATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAA
 AGTTTAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTACTACTGTTATAAATCATTAT
 TAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTTCG
 AACATTTGAGAAAATTTTGTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGAG
 GGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAAGTTGTACAAAAGTTGTACCAAATAGTTGTACAA
 ATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAATAAATAAGGATGACGC
 ATTAGAGAGATGTACCATTAGAGAATTTTTGGCAAGTCATTAAGAAAGAAAGAATAAATTATTTTTAAAAAT
 AAAAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTAT
 TAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCCTATATATTGCCCATAGAGTCA
 GTTAACCTATTTTTATATTTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAA
 AACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATC
 CAATCCAACCAATCACACAATCCTGATGAGATAACCCACTTTAAGCCACGCATCTGTGGCACATCTA
 CATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCT
 TCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTA
 ATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGAACAAGC
 TAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCCTTCAAACCTTCTGACGAAAGTCC
 GAGTTGGACTGAGTGGCGGCTACATAACGATGAGACGAATTCGAATCAAGATAATCCCCTTGGTTTCAA
 GGAAAGCTGGGGTTTCGGGAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTCACT
 GCACAGAGTCCTTGGATCTTGGACGGGAGATTGCGTTAACTATGCAGCATCTCGATTTTTCGGTTTCGA
 CCAGATCGGATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAAC
 TCTTCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGT
 GGAAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAA
 AATGCTTCTTCGTCTCCTATTTATAATATGGTTTGTATTGTTAATTTTGTCTTGTAGAAGAGCTTAATTA
 ATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTGGTGTAAATGTAATTCATTTACATAAGTGGAG
 TCAGAATCAGAATGTTTCCCTCCATAACTAAGTACTGACATGAAGACCTGCCGCTACAATTGTCTTATTTT
 GAACAACTAAAATTGAACATCTTTGCCACAACCTTTAAGTGGTTAATATAGCTCAAATATATGGTCAAG
 TTCAATAGATTATAATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAATACTACTA
 ATTTTATATCATCCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTG
 TCGTTTCCCGCCTTCAGTTTGCAAGCTGCTCTAGCCGTTAGCCAATACGCAAAACCGCCTCTCCCCGC
 GCGTTGGGAATTACTAGCGCGTGTGACAAAGCTGCATGCCGGTCAACATGGTGGAGCACGACACACT
 TGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAAGGGCAATTGAGACTTTTCAACAAAGG
 GTAATATCCGGAACCTCCTCGGATTCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAA
 AGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCC
 GACAGTGGTCCCAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAGAAAGAACGTTCCAACCA
 CGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATA
 TCAAAGATACAGTCTCAGAAGACCAAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAACCC
 TCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAGGAAGGTGGCTCCT
 ACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAA
 GATGGACCCCCACCCACGAGGAGCATCGTGGAAGAAAGAACGTTCCAACCACGTCTTCAAAGCAAGT
 GGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTC
 CTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAATCTTAATAGTTTTGATAAAGCGAACGT
 GGGGAAACCCGAACCAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTTGTCT
 TTCTTGCGTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACCAAGTACAACGTTTTCTTCACTGA
 AGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACGTGTACTTGTCTATTC
 TTGTGGTGTGGTCTTGGGAAAGAAAGCTTGTGGAGGCTGCTGTTAGCCCCATACATTACTTGTTA
 CGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTT
 CTTTCTTCTTCTTCTTGTGATTGGTCTATAAGAAATCTAGTATTTTCTTTGAAACAGAGTTTTCCCGTG
 GTTTTGAACCTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCGGGGCCCATGGCGAA
 AAACGTTGCGATTTTCGGCTTATTGTTTTCTTCTTGTGTTGGTTTCTTCTCAGATCTTCGCCTGCAGG
 CTCCTCAGCCAAAACGACACCCCATCTGTCTATCCACTGGCCCTGGATCTGCTGCCCAAACCTAACTC
 CATGGTGACCTGGGATGCCTGGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACCTCTG
 GATCCCTGTCCAGCGGTGTGCACACCTTCCAGCTGTCTGCGAGTCTGACCTCTACACTCTGAGCAGC
 TCAGTGACTGTCCCCTCCAGCACCTGGCCCAGCGAGACCGTCACCTGCAACGTTGCCCAACCGGCCA

GCAGCACCAAGGTGGACAAGAAAATTGTGCCAGGGATTGTGGTTGTAAGCCTTG CATATGTACAGTC
 CCAGAAGTATCATCTGTCTTCATCTTCCCCCAAAGGCCAAGGATGTGCTCACCATTACTCTGACTCCTA
 AGGTCACGTGTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTAGAT
 GATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGAGGAGCAATTCAACAGCACTTCCGCTCAG
 TCAGTGAACCTCCCATCATGCACCAGGACTGGCTCAATGGCAAGGAGCGATCGCTCACCATCACCATC
 ACCATCACCATCACCATTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTAT
 GTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGT
 TTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAATAAAAAAAAAAAAAAAAAAGAC
 CGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATT
 GAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTA
 ACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATAC
 GCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTCTATGTTACTA
 GATCTCTAGAGTCTCAAGCTTGGCGCGCCACGTGACTAGTGGCACTGGCCGTCGTTTTACAACGTCG
 TGA CTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGC
 GTAATAGCGAAGAGGCCCGCACC GATCGCCCTTCCCAACAGTTGCCGAGCCTGAATGGCGAATGCTA
 GAGCAGCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAACTATCAGTGTTTGACAGGAT
 ATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTTTA

Fig. 8G

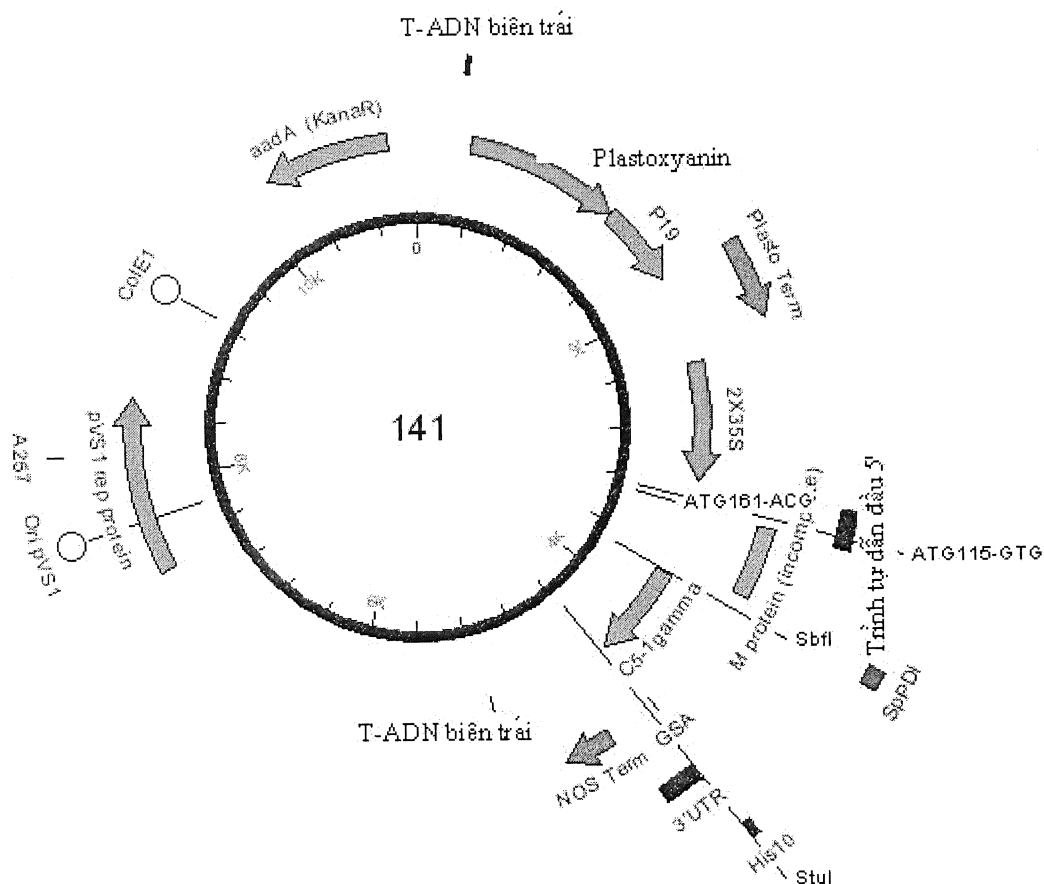


Fig. 8H, SEQ ID Sô:37

Caset biểu hiện số 1074 từ trình tự khởi động 2X35S đến trình tự kết thúc NOS. PDISP-Rab G-A/
Indonesia/5/2005 H5 TM+CT được gạch dưới

GTCAACATTTGGTGGAGACGACACACACTTGTCTACTCTCCAAATAATATCAAGAGATACAGTCTCAGAAGACCAACCA
AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAACCTCCTCGATTCCATTGCCAGCTATC
TGTCACCTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGA
AAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCA
TCGTGGAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGC
ACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTT
TCAACAAAGGGTAATATCCGGAACCTCCTCGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAA
GATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAG
ATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGAAAAAGAAGA
CGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACA
ATCCCCTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAAT
CTTAATAGTTTTTGATAAAAGCGAACGTGGGGAACCCGAACCAAACCTTCTTCTAACTCTCTCTCATC
TCTCTAAAGCAAACCTTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCG
GCACCAAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAGATCTTTTGTGGACACGTAGTGCGGCGC
CATTAATAACGTGTACTTGTCTATTCTTGTGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCT
GCTGTTGAGCCCCATACATTACTTGTTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTG
CTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAGTA
TTTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTG
CCCAAATTTGTGGGCCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTTCTTGTGTTGG
TTCTTCTCAGATCTTCGCCAAATTCCTATTTACACGATACCAGACAAGCTTGGTCCCTGGAGCCCCGA
TTGACATACATCACCTCAGCTGCCCAAACAATTTGGTAGTGGAGGACGAAGGATGCACCAACCTGTCA
GGGTTCTCCTACATGGAACCTAAAGTTGGATACATCTTAGCCATAAAAAATGAACGGGTTCACTTGCACA
GGCGTTGTGACGGAGGCTGAAACCTACACTAACTTCGTTGGTTATGTCACAACCACGTTCAAAAGAAAG
CATTTCCGCCCCAACACCAGATGCATGTAGAGCCGCGTACAACCTGGAAGATGGCCGGTGACCCAGATA
TGAAGAGTCTCTACACAATCCGTACCCTGACTACCGCTGGCTTCGAACTGTAAAAACCACCAAGGAGTC
TCTCGTTATCATATCTCCAAGTGTGGCAGATTTGGACCCATATGACAGATCCCTTCACTCGAGGGTCTT
CCCTAGCGGGAAGTGCTCAGGAGTAGCGGTGTCTTCTACCTACTGCTCCACTAACCACGATTACACCA
TTTGGATGCCCCGAGAATCCGAGACTAGGGATGTCTTGTGACATTTTACCAATAGTAGAGGGAAGAGAG
CATCCAAAGGGAGTGAGACTTGGCGCTTTGTAGATGAAAGAGGCCATATAAGTCTTTAAAGGAGCAT
GCAAACCTCAAGTTATGTGGAGTTCTAGGACTTAGACTTATGGATGGAACATGGGTGCGGATGCAAACAT
CAAATGAAACCAAATGGTGCCCTCCCGATCAGTTGGTGAACCTGCACGACTTTCGCTCAGACGAAATTG
AGCACCTTGTGTAGAGGAGTTGGTCAGGAAGAGAGAGGAGTGTCTGGATGCACTAGAGTCCATCATG
ACAACCAAGTCAGTGAGTTTCAGACGTCTCAGTCATTTAAGAAAACTTGTCCTGGGTTTGGAAAAGCA
TATACCATATTCAACAAGACCTTGATGGAAGCCGATGCTCACTACAAGTCAGTCAGAACCTTGGAAATGAG
ATCCTCCCTTCAAAAGGGTGTTTAAGAGTTGGGGGAGGTGTCACTCCTCATGTGAACGGGGTGTTTTTTC
AATGGTATAATATTAGGACCTGACGGCAATGTCTTAATCCAGACATGATGCAATCATCCCTCCTCCAGCAA
CATATGGAGTTGTTGGAATCCTCGGTTATCCCCCTGTGACACCCCTGGCAGACCCGTCTACCGTTTTTC
AAGGACGGTGACGAGGCTGAGGATTTTGTGAAGTTACCTTCCCGATGTGCACAATCAGGTCTCAGG
AGTTGACTTGGGTCTCCCGAAGTGGGGGAAGTACAAATACTGTCAATTTATTCAACAGTGGCGAGTTC
CCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGATCGTTACAATGCAG
AATTTGCATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTATTCGGTGTGCATTTCTATGTTTGGTG
AGCGTTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGG
TCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATAAAAAAGACCGGGAATT
CGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTG
TTGCCGGTCTTGCAGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAA
TGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAG
AAAACAAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGAT

Fig. 8I, SEQ ID Số:38

Trình tự axit amin của PDISP-Rab G-A/Indonesia/5/2005 H5 TM+CT

MAKNVAIFGLLFSLVLVPSQIFAKFPIYTIPDKLGPWSPIDIHHLSCPNNLVEDEGCTNL
SGFSYMEKLVGYILAIKMNGFTCTGVVTEAETYTNFVGYVTTTTFKRKHFRPTPDACRAA
YNWKMGADPRYEESLHNPYPDYRWLRTVKTTKESLVIISPSVADLDPYDRSLHSRVFP
SGKCSGVAVSSTYCSTNHDYTIWMPENPRLGMSCDIFTNSRGKRASKGSETCGFVDE
RGLYKSLKGACKLKLGVGLRLMDGTWVAMQTSNETKWCPPDQLVNLHDFRSDEIE
HLVVEELVRKREECLDALESIMTTKSVSFRRLSHLRKLVPGFGKAYTIFNKTLMEDAHAH
KSVRTWNEILPSKGCLRVGGGRCHPHVNGVFFNGIILGPDGNVLIPEMQSSLLQQHMELL
ESSVIPLVHPLADPSTVFKDGEAEDFVEVHLPDVHNQVSGVDLGLPNWGKYQILSIYS
TVASSLALAIMMAGLSLWMCSNGSLQCRICI*

Fig. 8J, SEQ ID Số:39

Sản phẩm thiết kế 144 từ t-ADN từ trái sang phải (được gạch dưới). Hệ thống biểu hiện 2X35S-CPMV-HT-PDISP-NOS thành BeYDV + replicaza với caset biểu hiện chất ức chế bất hoạt Plastoxyanin-P19-Plastoxyanin

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTTTTA
ATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAA
AGTTTAAGTTAGCAAGTGTGTACATTTTTACTTGAACAAAAATATTCACCTACTACTGTTATAAATCATTAT
TAAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTTGC
AACATTTGAGAAAAATTTGTTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGAG
GGAGAATAAAAAACATAATGTGAGTAGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTACAA
ATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGACGC
ATTAGAGAGATGTACCATTAGAGAATTTTTGGCAAGTCATTAAGAAAGAAAGAAATTAATTTTTAAAT
AAAAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGTTGTAT
TAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCTATATATTGCCCCATAGAGTCA
GTTAACTCATTTTTATATTTATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAAA
AACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATAACATC
CAATCCAACCAATCACAACAATCCTGATGAGATAACCCACTTTAAGCCCACGCATCTGTGGCACATCTA
CATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCACCCAT
TCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAGACTA
ATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGAACAAGC
TAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCCTTCAAACCTTCTGACGAAAGTCC
GAGTTGGACTGAGTGGCGGCTACATAACGATGAGACGAATTCGAATCAAGATAATCCCCTTGGTTTCAA
GGAAAGCTGGGTTTTCGGGAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCACT
GCACAGAGTCCTTGGATCTTGGACGGGAGATTCGGTTAACTATGCAGCATCTCGATTTTTCGGTTTGA
CCAGATCGGATGTACCTATAGTATTCGGTTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAAC
TCTTCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGT
GGAAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAA
AATGCTTCTTCGTCTCCTATTTATAATATGGTTTGTTATTGTTAATTTGTTCTTGTAGAAGAGCTTAATTA
ATCGTTGTTGTTATGAAATACTATTTGTATGAGATGAAGTGGTGAATGTAATTCATTTACATAAGTGGAG
TCAGAATCAGAATGTTTCCTCCATAACTAACTAGACATGAAGACCTGCCGCTACAATTGTCTTATATTT
GAACAATAAATTGAACATCTTTGCCACAACCTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAG
TTCAATAGATTAATAATGGAATATCAGTTATCGAAATTCATTAACAATCACTTAACGTTATTAACCTACTA
ATTTTATATCATCCCCTTTGATAAATGATAGTACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTG

TCGTTTCCCGCCTTCAGTTTTCGAAGCTGCTCTAGCGGTGTAGCCAAATACGCAAAACCGCCTCTCCCGCGCGTTGGGAATTACTAGCGCGTGTGACACGCGTGGCGCGCCCTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACCTCTATCTTATAACCGGCGTGGAGGCATGGAGGCAAGGGCATTTTTGTTAAATTAAGTAGTTAGTGGAAAATGACGTCACTTACTTAAAGACGAAGTCTTGCACAAAGGGGGGCCACGCGAATTTTAATATTACCGGCGTGGCCCCACCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAAGTTCCCGCCCACTAGGGTTAAAGGTGTTACACTATAAAAGCATATACGATGTGATGGTATTTGATAAAAGCGTATATTGTATCAGGTATTTCCGTGCGATACGAATTATTCGTACAAGCTTCTTAAGCCGGTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAAGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTCAACAAAAGGTAATATCCGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACATATCCCACTATCCTTGCAGAACCCCTTCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAACCCGAACCAAACCTTCTTCTAACTCTCTCTCATCTCTCTAAAGCAAACCTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTGAGATCGTGCTTCGGCACACGTAACCGTTTTCTTCACTGAAGCGAAATCAAAGATCTTTTGTGGACACGTAGTGCGGCGCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTCTTGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTGAGCCCATACATTACTTGTTACGATTCTGCTGACTTTGCGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAGTATTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCAAATTTGTGCGGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTCTTGTGTTGGTTCTTCTCAGATCTTCGCCTGCAGGCTCCTCAGCCAAAACGACACCCCATCTGTCTATCCACTGGCCCTTGATCTGCTGCCAAACTAACTCCATGGTGACCCTGGGATGCCTGGTCAAGGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACCTGGATCCCTGTCCAGCGGTGTGCACACCTTCCAGCTGTCTGCAGTCTGACCTCTACACTCTGAGCAGCTCAGTGACTGTCCCTCCAGCACCTGGCCCAGCGAGACCGTCACTGCAACGTTGCCACCCGCGCCAGCAGCACCAAGGTGGACAAGAAAATTGTGCCCAGGGATTGTGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCTTCCCCCAAAGCCCAAGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACATCAGCAAGGATGATCCCGAGGTCCAGTTCAGCTGGTTTGTAGATGATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGGAGGAGCAGTTCAACAGCACTTTCCGCTCAGTCAGTGAACCTCCCATCATGCACCAGGACTGGCTCAATGGCAAGGAAGGCCTATTTTCTTAGTTTGAATTTACTGTTATTGCGGTGTGCATTTCTATGTTTGGTGAGCGTTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTAATTTCTTTGTGAGCTCCTGTTAGCAGGTGCTCCCTCAGCAAGGACACAAAAAGATTTTAAATTTATTAACCAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCGGTGTCTATCATATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGGGGTACCGAGCTCGAATTCGAGTGTACTTCAAGTCAGTTGGAAATCAATAAAATGATTATTTTATGAATATATTTCTTGTGCAAGTAGATAGAAATTACATATGTTACATAACACACGAAATAAACAAACCAACACAATCCAAACAAACACCCCAACAAATAACTATATATATCCTCGTATGAGGAGAGGCACGTTTCACTGACTCGACGATTCCCGAGCAAAAAAGTCTCCCGTCACACATATAGTGGGTGACGCAATTA TCTTCAAAGTAATCCTTCTGTTGACTTGTCTTGTGATAACATCCAGTCTTCGTGAGGATTGCAAGAATTATAGAAGGGATCCCACCTTTATTTCTTCTTTTTCCATATTTAGGGTTGACAGTGAAATCAGACTGGCAACCTATTAAATTGCTTCCACAATGGGACGAACCTGAAGGGGATGTGTCGATGATATTATAGGTGGCGTGTTCATCGTAGTTGGTGAAGTCGATGGTCCCGTTCAGTAGTTGTGTGCGCCGAGACTTCTAGCCCAGGTGGTCTTTCCGGTACGAGTTGGTCCGCAGATGTAGAGGCTGGGGTGTCTGACCCAGTCTTCCCTCATCCTGGTTAGATCGGCCATCCACTCAAGGTCAGATTGTGCTTGATCGTAGGAGACAGGATGTATGAAAGTGAGGCATCGATGCTTACATGATATAGGTGCGTCTCTCTCCAGTTGTGCAGATCTTCGTGGCAGCGGATCTGATTCTGTGAAGGGCGACACGTAAGTCTCAGGTTGTGGAGGAATAATTTGTTGGCTGAATATT

CCAGCCATTGAAGCTTTGTTGCCCATTCATGAGGGAATTCTTCTTTGATCATGTCAAGATACTCCTCCTT
 AGACGTTGCAGTCTGGATAATAGTTCGCCATCGTGCGTCAGATTTGCGAGGAGAGACCTTATGATCTCG
 GAAATCTCCTCTGGTTTAAATATCTCCGTCCTTTGATATGTAATCAAGGACTTGTTTAGAGTTTCTAGCTG
 GCTGGATATTAGGGTGATTTCTTCAAAATCGAAAAAGAAGGATCCCTAATAACAAGGTTTTTATCAAG
 CTGGATAAGAGCATGATAGTGGGTAGTGCCATCTTGATGAAGCTCAGAAGCAACACCAAGGAAGAAAA
 TAAGAAAAGGTGTGAGTTTCTCCAGAGAACTGGAATAAATCATCTCTTTGAGATGAGCACTTGGGGT
 AGGTAAGGAAAACATATTTAGATTGGAGTCTGAAGTTCTTGCTAGCAGAAGGCATGTTGTTGTGACTCC
 GAGGGGTTGCCTCAAACCTATCTTATAACCGGCGTG GAGGCATGGAGGCAAGGGCATTITGGTAATT
 TAAGTAGTTAGTGGAAAATGACGTCATTTACTTAAAGACGAAGTCTTGCGACAAGGGGGGCCACGCC
 GAATTTTAATATTACCGGCGTGGCCCCACCTTATCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAA
 GTAGAAAAGTTCCCGCCCACTAGGGTTAAAGGTGTTTACACTATAAAAGCATATACGATGTGATGGTAT
 TTGATGGAGCGTATATTGTATCAGGTATTTCCGTCGGATACGAATTATTCGTACGGCCGGCCACTAGTG
 GCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGCA
 GCACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTT
 GCGCAGCCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGATTGTGTTTTCCCGCCTTCAGT
 TTAAACTATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTTTA

Fig 8K

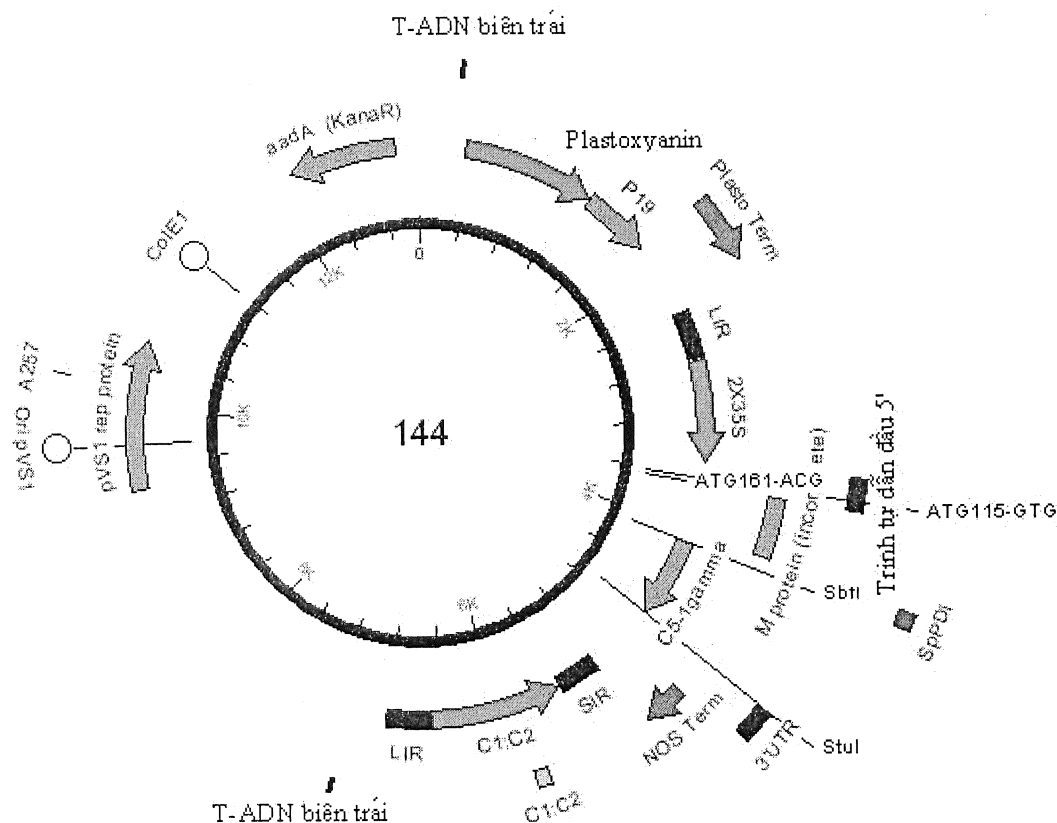


Fig. 8L, SEQ ID Số:40

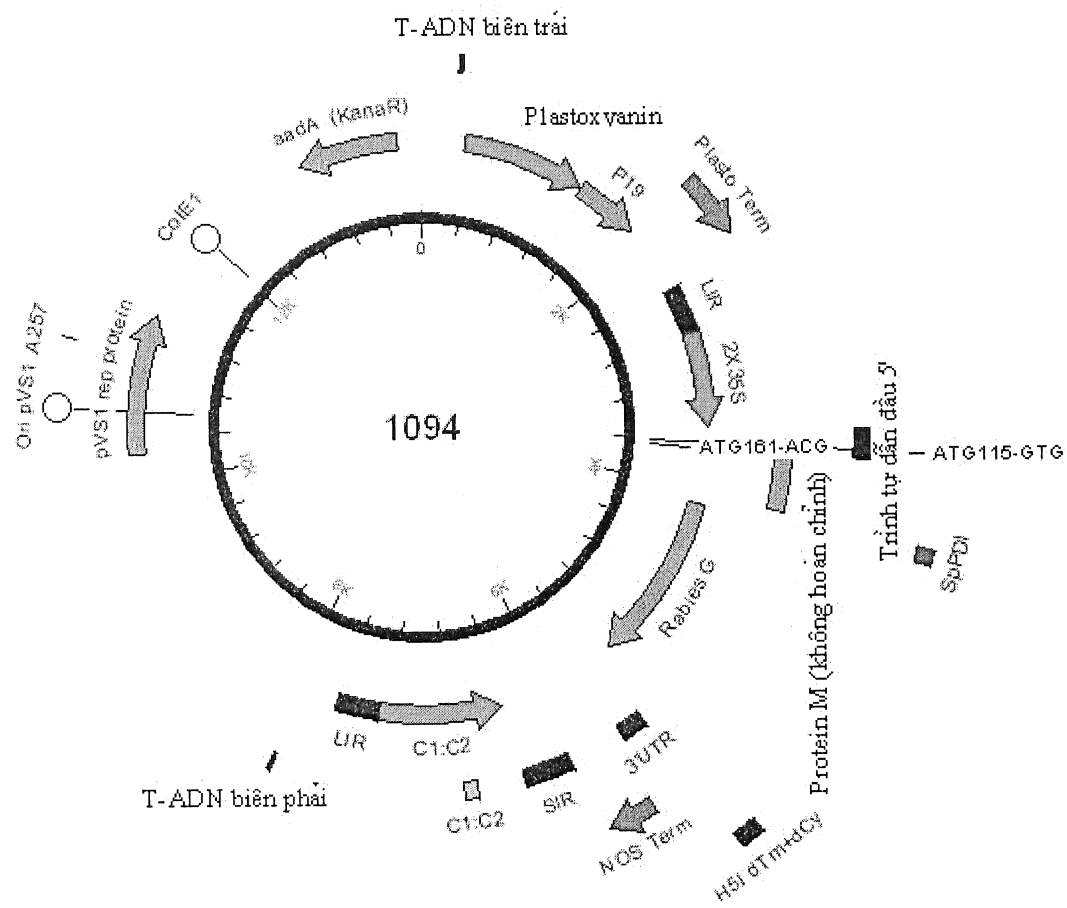
Caset biểu hiện số 1094 từ BeYDV LIR từ phải sang trái. PDISP-Rab G-A/Indonesia/5/2005 H5 TM+CT được gạch dưới

CTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACCTCTATCTTATAACCGGCGTGAG
GCATGGAGGCAAGGGCATTGTTGTAATTTAAGTAGTTAGTGGAAAATGACGTCATTTACTTAAAGACGA
AGTCTTGCGACAAGGGGGGGCCACGCCGAATTTTAATATTACCGGCGTGGCCCCACCTTATCGCGAGT
GCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAAGTTCCCGCCCACTAGGGTTAAAGGTGTTACAC
TATAAAGCATATACGATGTGATGGTATTTGATAAAGCGTATATTGTATCAGGTATTTCCGTCCGATACG
AATTATTCGTACAAGCTTCTTAAGCCGGTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATA
TCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACC
TCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCT
ACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAA
GATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGT
GGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTC
AGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCCTCCTCGGATTCCATTG
CCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTG
CGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAGATGGACCCCCACCCA
CGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCT
CCTACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTCCTCTATATAAGGAAGTT
CATTTTATTTGGAGAGGTATTTAAATCTTAATAGTTTTGATAAAAGCGAACGTGGGGAAACCCGAACCA
AACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAGCAAACCTTCTCTTGTCTTTCTTGCCTGAGCGATC
TTCAACGTTGTGAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAGATCT
CTTTGTGGACACGTAGTGGCGGCCATTAAATAACGTGTACTTGTCTATTCTTGTGGTGTGGTCTTG
GGAAAAGAAAGCTTGCTGGAGGCTGCTGTTGAGCCCCATACATTACTTGTACGATTCTGCTGACTTTC
GGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTCTTCTTCTTCTGC
TGATTGGTTCTATAAGAAATCTAGTATTTCTTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAA
AGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTGCGGGCCCATGGCGAAAAACGTTGCGATTTTCG
GCTTATTGTTTTCTTCTTGTGTTGGTTCCCTTCTCAGATCTTCGCCAAATTCCTATTTACACGATACCA
GACAAGCTTGGTCCCTGGAGCCCGATTGACATACATCACCTCAGCTGCCCAAACAATTTGGTAGTGGA
GGACGAAGGATGCACCAACCTGTGAGGGTTCTCTACATGGAACCTTAAAGTTGGATACATCTTAGCCAT
AAAAATGAACGGGTTCACTTGACAGGCGTTGTGACGGAGGCTGAAACCTACACTAAGTTGGTTGTTA
TGTCACAACCACGTTCAAAGAAAGCATTTCGCCCAACACCAGATGCATGTAGAGCCGCGTACAACCTG
GAAGATGGCCGGTGACCCAGATATGAAGAGTCTCTACACAATCCGTACCCTGACTACCGCTGGCTTC
GAACTGTAAAACCAAGGAGTCTCTCGTTATCATATCTCCAAGTGTGGCAGATTTGGACCCATATG
ACAGATCCCTTCACTCGAGGGTCTTCCCTAGCGGGAAGTGCTCAGGAGTAGCGGTGTCTTCTACCTAC
TGCTCCACTAACCACGATTACACCATTTGGATGCCGAGAATCCGAGACTAGGGATGTCTTGTGACATT
TTTACCAATAGTAGAGGGAAGAGAGCATCCAAAGGGAGTGAGACTTGCGGCTTTGTAGATGAAAGAGG
CCTATATAAGTCTTTAAAGGAGCATGCAAACCTCAAGTTATGTGGAGTTCTAGGACTTAGACTTATGGAT
GGAACATGGGTCGCGATGCAAACATCAAATGAAACCAAATGGTGCCCTCCCGATCAGTTGGTGAACCT
GCACGACTTTGCTCAGACGAAATTGAGCACCTTGTGTAGAGGAGTTGGTCAGGAAGAGAGAGGAGT
GTCTGGATGCACTAGAGTCCATCATGACAACCAAGTCAGTGAGTTTCAGACGCTCTCAGTCATTTAAGAA
AACTTGTCCCTGGGTTTGGAAAAGCATATACCATATTCAACAAGACCTTGATGGAAGCCGATGCTCACT
ACAAGTCAGTCAGAACTTGAATGAGATCCTCCCTTCAAAGGGTGTTTAAAGAGTTGGGGGGAGGTGT
CATCCTCATGTGAACGGGGTGTGTTTCAATGGTATAATATTAGGACCTGACGGCAATGTCTTAATCCCAG
AGATGCAATCATCCCTCCTCCAGCAACATATGGAGTTGTTGGAATCCTCGGTTATCCCCCTTGTGCACC
CCCTGGCAGACCCGTCTACCGTTTTCAAGGACGGTGACGAGGCTGAGGATTTTGTGAAGTTACCTT
CCCGATGTGCACAATCAGGTCTCAGGAGTTGACTTGGTCTCCCGAACTGGGGGAAGTATCAAATACT
GTCAATTTATTCAACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGAT
GTGCTCCAATGGATCGTTACAATGCAGAATTTGATTTAAAGGCCTATTTCTTAGTTTGAATTTACTGT
TATTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTATGTAATTTA

ATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAA
AAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGC
AATAAAGTTTCTTAAGATTGAATCCTGTTGCCGCTTTCGATGATTATCATATAATTTCTGTTGAATTAC
GTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCC
GCAATTATACATTTAATACGCGATAGAAAAACAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGC
GGTGTCACTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGGGGTACCGAGCTCGAATTCCGA
GTGTACTTCAAGTCAGTTGGAAATCAATAAAATGATTATTTTATGAATATATTTTCATTGTGCAAGTAGATA
GAAATTACATATGTTACATAACACACGAAATAAACAAAAAACACAATCCAAAAACAAACACCCCAACAA
AATAACACTATATATATCCTCGTATGAGGAGAGGCACGTTCACTGACTCGACGATTCCCGAGCAAAAA
AGTCTCCCCGTACACATATAGTGGGTGACGCAATTATCTTCAAAGTAATCCTTCTGTTGACTTGTCAAT
GATAACATCCAGTCTTCGTGAGGATTGCAAAGAATTATAGAAGGGATCCCACCTTTTATTTCTTCTTTTT
TCCATATTTAGGGTTGACAGTGAAATCAGACTGGCAACCTATTATTGCTTCCACAATGGGACGAACCT
GAAGGGGATGTCGTGATGATATTATAGGTGGCGTGTTTCATCGTAGTTGGTGAAGTCGATGGTCCCGT
TCCAGTAGTTGTGTCGCCCGAGACTTCTAGCCCAGGTGGTCTTTCCGGTACGAGTTGGTCCGCAGATG
TAGAGGCTGGGGTGTCTGACCCAGTCCTTCCCTCATCCTGGTTAGATCGGCCATCCACTCAAGGTCA
GATTGTGCTTGATCGTAGGAGACAGGATGTATGAAAGTGTAGGCATCGATGCTTACATGATATAGGTGC
GTCTCTCTCCAGTTGTGCAGATCTTCGTGGCAGCGGAGATCTGATTCTGTGAAGGGCGACACGTAAGT
CTCAGGTTGTGGAGGAAATAATTTGTTGGCTGAATATTCCAGCCATTGAAGCTTTGTTGCCCATTCATGA
GGGAATTCCTCTTTGATCATGTCAAGATACTCCTCCTTAGACGTTGCAGTCTGGATAAATAGTTCGCCATC
GTGCGTCAGATTTGCGAGGAGAGACCTTATGATCTCGGAAATCTCCTCTGTTTTAATATCTCCGTCCCTT
TGATATGTAATCAAGGACTTGTTTAGAGTTTCTAGCTGGCTGGATATTAGGGTGATTTCTTCAAATCG
AAAAAGAAGGATCCCTAATACAAGGTTTTTTATCAAGCTGGATAAAGAGCATGATAGTGGGTAGTGCCA
TCTTGATGAAGCTCAGAAGCAACACCAAGGAAGAAAATAAGAAAAGGTGTGAGTTTCTCCCAGAGAAAC
TGGAATAAATCATCTCTTTGAGATGAGCACTTGGGGTAGGTAAAGGAAAACATATTTAGATTGGAGTCTG
AAGTTCTTGCTAGCAGAAGGCATGTTGTTGTGACTCCGAGGGGTTGCCTCAAACCTCTATCTTATAACCG
GCGTGGAGGCATGGAGGCAAGGGCATTGTTGTTAAGTAGTTAGTGGAATGACGTCATTTACTT
AAAGACGAAGTCTTGCGACAAGGGGGGGCCACGCCGAATTTAATATTACCGGCGTGCCCCACCTTA
TCGCGAGTGCTTTAGCACGAGCGGTCCAGATTTAAAGTAGAAAAGTTCCCGCCCACTAGGGTTAAAGG
TGTTCACTATAAAAGCATATACGATGTGATGGTATTTGATGGAGCGTATATTGTATCAGGTATTTCCG
TCGGATACGAATTATTCGTAC

Fig. 8M

Sự thể hiện dưới dạng sơ đồ sản phẩm thiết kế số 1094



G bệnh đại

Western blot trên các sinh khối

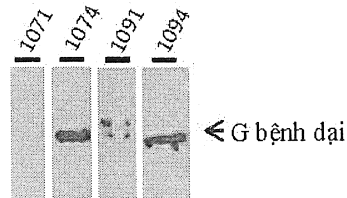


Fig. 8N

Phương pháp sắc ký loại trừ kích thước cho dịch chiết được lọc và cô đặc

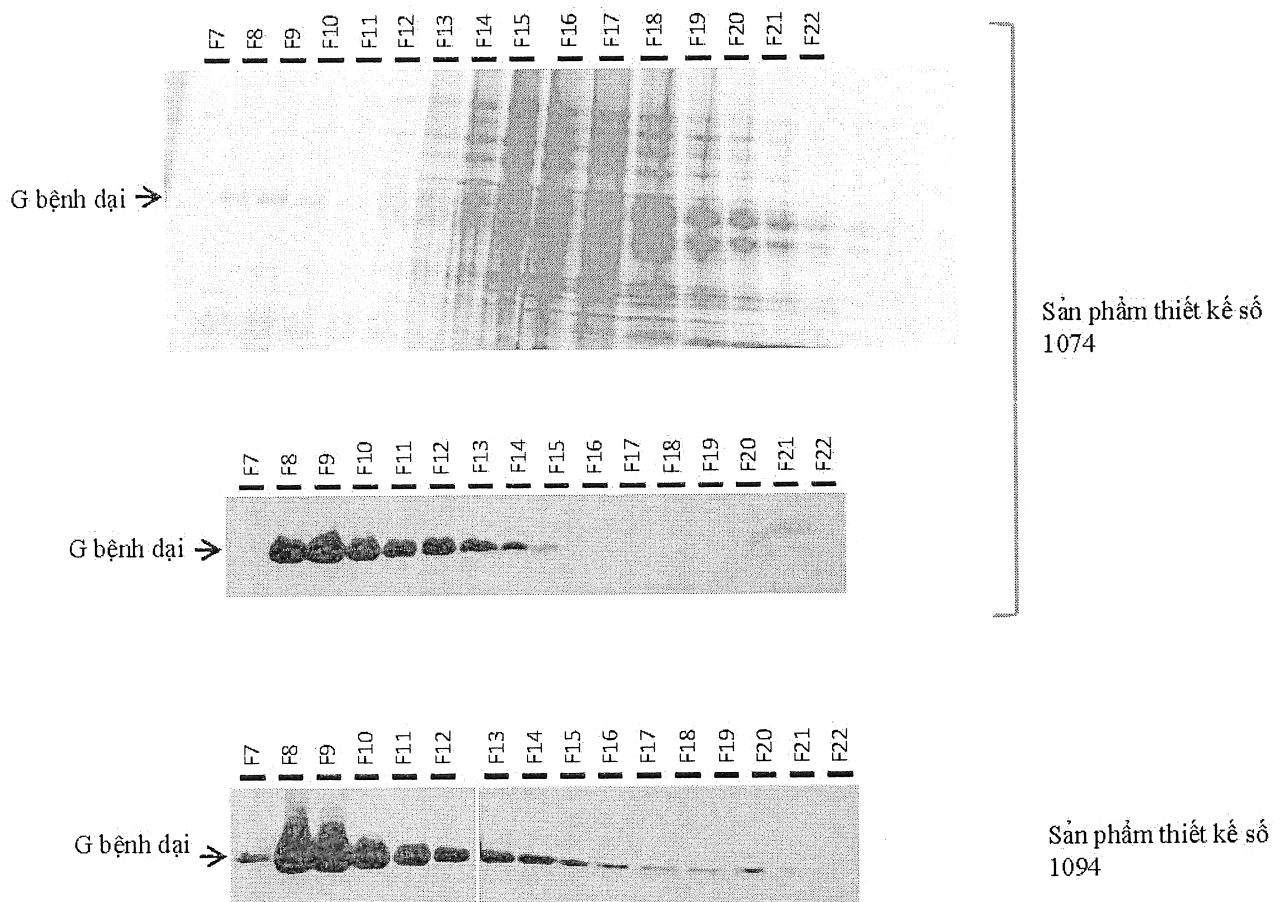


Fig. 8O

Kháng thể sơ cấp: Santa Cruz SE-57995 0,25 $\mu\text{g}/\mu\text{L}$

Kháng thể thứ cấp: Chuột kháng dê, JIR, 115-035-146 (1/10000)

G bệnh đại

Tinh chế (1074)

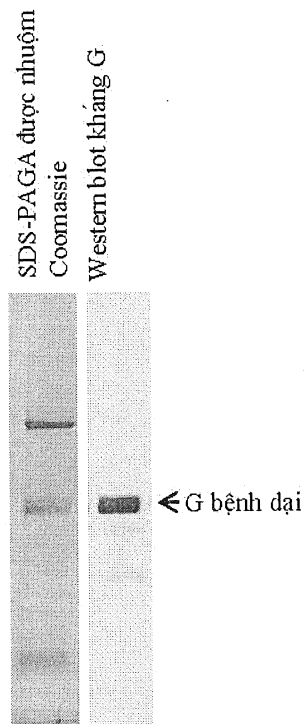


Fig. 9A

Kính hiển vi truyền điện tử trên VLP được tinh chế (1074) (thanh là 100nm)

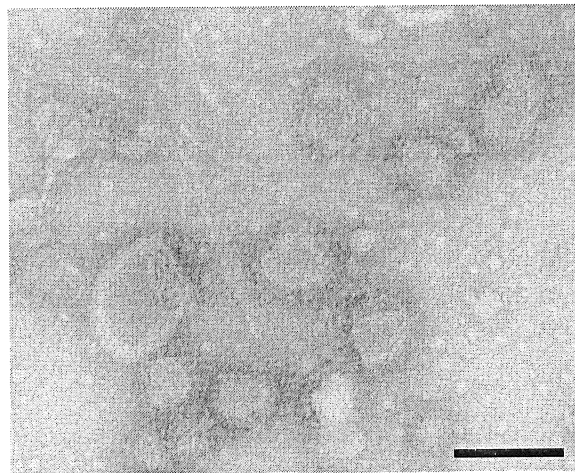


Fig. 9B

Fig. 10A, SEQ ID Số: 20

IF-wtSp-VZVgE.c

TGCCCAAATTTGTCTGGGCCCATGGGGACAGTTAATAAACCTGTGGTG

Fig. 10B, SEQ ID Số: 21

IF-H5dTm+VZVgE.r

AATTGACAGTATTTGTCTAGAAAGTGGTGACGTTCCGGGGTTTACG

Fig. 10C, SEQ ID Số: 22

gen VZV gE được tổng hợp (tương ứng với nt 3477-5348 từ Số gia nhập ngân hàng gen AY013752.1 (peptit tín hiệu bẩm sinh được in đậm, vùng xuyên màng bẩm sinh và đuôi chất nguyên sinh được gạch dưới))

ATGGGGACAGTTAATAAACCTGTGGTGGGGGTATTGATGGGGTTCGGAATTATCACGGGAACGTTGC
 GTATAACGAATCCGGTCAGAGCATCCGTCTTGCATACGATGATTTTCACACCGATGAAGACAACTG
 GATACAACTCCGTATATGAGCCTTACTACCATTGAGATCATGCGGAGTCTTCATGGGTAAATCGGGGA
 GAGTCTTCGCGAAAAGCGTACGATCATAACTCACCTTATATATGGCCACGTAATGATTATGATGGATTTT
 TAGAGAACGCACACGAACACCATGGGGTGTATAATCAGGGCCGTGGTATCGATAGCGGGGAACGGTTA
 ATGCAACCCACACAAATGTCTGCACAGGAGGATCTTGGGGACGATACGGGCATCCACGTTATCCCTAC
 GTTAAACGGCGATGACAGACATAAAATTGTAAATGTGGACCAACGTCAATACGGTGACGTGTTTAAAGG
 AGATCTTAATCCAAAACCCCAAGGCCAAAGACTCATTGAGGTGTCAGTGGAAGAAAATCACCCGTTTAC
 TTTACGCGCACCGATTGAGCGGATTTATGGAGTCCGGTACACCGAGACTTGGAGCTTTTTGCCGTCATT
 AACCTGTACGGGAGACGCGAGCGCCCGCCATCCAGCATATATGTTTAAAACATACAACATGCTTTCAAGA
 CGTGGTGGTGGATGTGGATTGCGCGGAAAATACTAAAGAGGATCAGTTGGCCGAAATCAGTTACCGTT
 TTCAAGGTAAGAAGGAAGCGGACCAACCGTGGATTGTTGTAACACGAGCACACTGTTTGATGAACTC
 GAATTAGACCCCCCGAGATTGAACCGGGTGTCTTGAAAGTACTTCGGACAGAAAAACAATACTTGGGT
 GTGTACATTTGGAACATGCGCGGCTCCGATGGTACGTCTACCTACGCCACGTTTTTGGTCACCTGGAAA
 GGGGATGAAAAACAAGAAACCCTACGCCCCGAGTAACCTCAACCAAGAGGGGCTGAGTTTCATAT
 GTGGAATTACCACTCGCATGTATTTTCAGTTGGTGATACGTTTAGCTTGGCAATGCATCTTCAGTATAAG
 ATACATGAAGCGCCATTTGATTTGCTGTTAGAGTGGTGTATGTCCCATCGATCCTACATGTCAACCAA
 TGCGGTTATATTCTACGTGTTTGTATCATCCCAACGCACCCCAATGCCTCTCTCATATGAATCCGGTTG
 TACATTTACCTCGCCACATTTAGCCCAGCGTGTGCAAGCACAGTGTATCAAAATTGTGAACATGCAGA
 TAACCTACACCGCATATTGTCTGGGAATATCTCATATGGAGCCTAGCTTTGGTCTAATCTTACACGACGG
 GGGCACCGTAAAGTTTGTAGATACACCCGAGAGTTTGTCTGGGATTATACGTTTTTGTGGTGTATTTT
 AACGGGCATGTTGAAGCCGTAGCATACACTGTTGTATCCACAGTAGATCATTTTGTAAACGCAATTGAA
 GAGCGTGGATTTCCGCCAACGGCCGTCAGCCACCGGCGACTACTAAACCCAAGGAAATTACCCCCG
 TAAACCCCGGAACGTCACCACTTCTACGATATGCCGCATGGACCGGAGGGCTTGCAGCAGTAGTACTT
 TTATGTCTCGTAATATTTTTAATCTGTACGGCTAAACGAATGAGGGTTAAAGCCTATAGGGTAGACAAGT
 CCCCCTATAACCAAGCATGTATTACGCTGGCCTTCCAGTGGACGATTTGAGGACTCGGAATCTACG
 GATACGGAAGAAGATTTGATAACGCGATTGGAGGGAGTCACGGGGGTTTCAGTTACACGGTGTATAT
 AGATAAGACCCGGTGA

Fig. 10D, SEQ ID Số: 23

VZVgE+H5dTm.c

GCCACTGTTGAATAAATTGACAGTATTTGTCGTAGAAGTGGTGACG

Fig. 10E, SEQ ID Số: 24

Caset biểu hiện số 946, từ Pacl (ngược chiều với trình tự khởi động) với Ascl (xuôi chiều ngay với trình tự kết thúc NOS). VZV gE-A/Indonesia/5/2005 H5 TM+CT được gạch dưới.

TTAATTAAGTCGACAAGCTTGCATGCCGGTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATA
TCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTC
CTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAA
ATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCCTCTGCCGACAGTGGTCCCAAAGATGGAC
CCCCACCCACGAGGAGCATCGTGGAAAAAGAACGTTCCAACCACGCTTCAAAGCAAGTGGATTGATGT
GATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAG
GGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGT
ACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCC
ATCGTTGAAGATGCCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAA
AGAAGACGTTCCAACCACGCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACG
CACAATCCCCTATCCTTCGCAAGACCCCTCCTCTATATAAGGAAGTTCATTTCACTTTGGAGAGGTATTAA
AATCTTAATAGTTTTTGATAAAAGCGAACGTGGGGAACCCGAACCAAACCTTCTTCTAACTCTCTCTCA
TCTCTCTTAAAGCAAACCTCTCTCTTGTCTTTCTTGCCTGAGCGATCTTCAACGTTGTCAGATCGTGCTTC
GGCACCAGTACAACGTTTTCTTCTACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGCGGCGCC
ATTAAATAACGTGTACTTGTCTTATTCTTGTCTGGTGTGGTCTTGGGAAAAGAAAGCTTGTGGAGGCTGCT
GTTACAGCCCCATACATTACTTGTACGATTCTGTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTG
ACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTGCTGATTGGTTCATAAGAAATCTAGTATTTT
CTTTGAAACAGAGTTTTTCCCGTGGTTTTTCAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCA
AATTTGTCTGGGCCCATTGGGACAGTTAATAAACCTGTGGTGGGGTATTGATGGGGTTCGGAATTATCACG
GGAACGTTGCGTATAACGAATCCGGTCAGAGCATCCGCTTTCGCGATACGATGATTTTACACCGATGAAGA
CAAACCTGGATACAACTCCGTATATGAGCCTTACTACCATTCAGATCATGCGGAGTCTTCATGGGTAAATC
GGGAGAGTCTTTCGCGAAAAGCGTACGATCATACTACCTTATATATGGCCACGTAATGATTATGATGGA
TTTTTAGAGAACGCACACGAACACCATGGGGTGTATAATCAGGGCCGTGGTATCGATAGCGGGGAACGGTT
AATGCAACCCACACAAATGTCTGCACAGGAGGATCTTGGGGACGATACGGGCATCCACGTTATCCCTACGT
TAAACGGCGATGACAGACATAAAATTGTAAATGTGGACCAACGTCAATACGGTGACGTGTTAAAGGAGAT
CTTAATCCAAAACCCCAAGGCCAAAGACTCATTGAGGTGTCAGTGGAAGAAAATCACCCGTTTACTTTACG
CGCACCGATTACGCGGATTTATGGAGTCCGGTACACCGAGACTTGGAGCTTTTTGCCGTCATTAACCTGTA
CGGGAGACGCAGCGCCCCGCCATCCAGCATATATGTTTAAACATACAACATGCTTTCAAGACGTGGTGGTG
GATGTGGATTGCGCGGAAAATACTAAAGAGGATCAGTTGGCCGAAATCAGTTACCGTTTTCAAGGTAAGAA
GGAAGCGGACCAACCGTGGATTGTTGTAAACACGAGCACACTGTTTGATGAACTCGAATTAGACCCCCCG
AGATTGAACCGGGTGTCTTGAAAGTACTTCGGACAGAAAAACAATACTTGGGTGTGTACATTTGGAACATG
CGCGGCTCCGATGGTACGTCTACCTACGCCAGTTTTTGGTCACTTGGAAAGGGGATGAAAAACAAGAAA
CCCTACGCCCGCAGTAACCTCCTCAACCAAGAGGGGCTGAGTTTCATATGTGGAATTACCACTCGCATGTAT
TTTCAGTTGGTGATACGTTTAGCTTGGCAATGCATCTTCAGTATAAGATACATGAAGCGCCATTTGATTTG
CTGTTAGAGTGGTTGTATGTCCCCATCGATCCTACATGTCAACCAATGCGGTTATATTCTACGTGTTTGTA
TCATCCCAACGCACCCCAATGCCTCTCTCATATGAATTCGGGTGTACATTTACCTCGCCACATTTAGCCC

AGCGTGTTGCAAGCACAGTGTATCAAAATTGTGAACATGCAGATAACTACACCGCATATTGTCTGGGAATA
 TCTCATATGGAGCCTAGCTTTGGTCTAATCTTACACGACGGGGGCACCACGTTAAAGTTTGTAGATACACC
 CGAGAGTTTGTCTGGGATTATACGTTTTTGTGGTGTATTTTAACGGGCATGTTGAAGCCGTAGCATACACTG
 TTGTATCCACAGTAGATCATTTTGTAAACGCAATTGAAGAGCGTGGATTTCGCGCAACGGCCGGTCAGCCA
 CCGGCGACTACTAAACCCAAGGAAATTACCCCCGTAAACCCCGGAACGTCACCACTTCTACGACAAATACT
 GTCAATTTATTCAACAGTGGCGAGTTCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGT
 GCTCCAATGGATCGTTACAATGCAGAATTTGCATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTA
 TTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAAT
 TTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAAAA
 AAAAAAAAAAAAAAGACCGGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAAT
 AAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTT
 AAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCA
 ATTATACATTTAATACGCGATAGAAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGT
 CATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGGCC

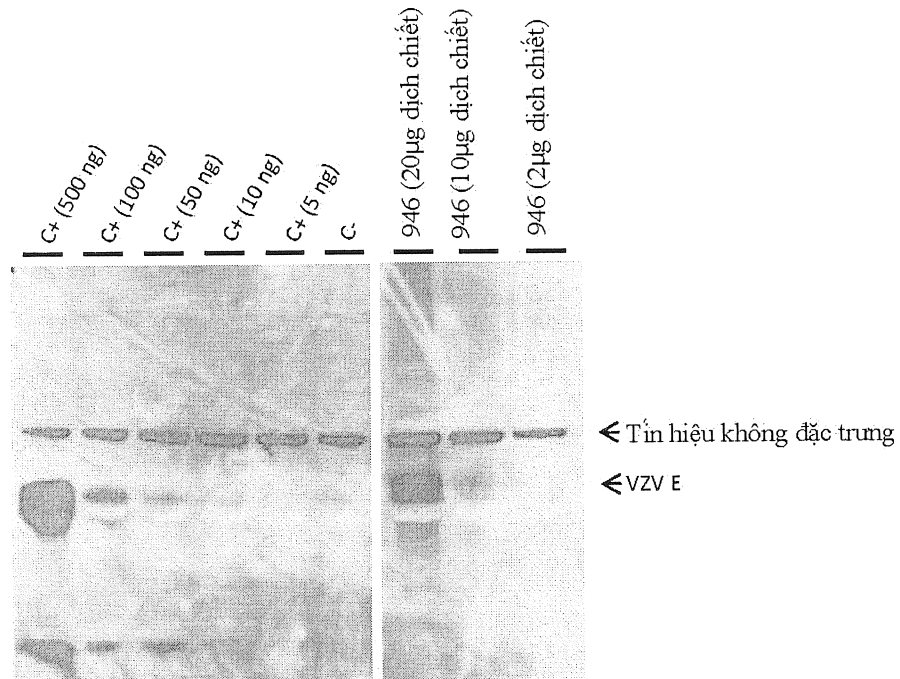
Fig. 10F, SEQ ID Số: 25

Trình tự axit amin của VZV gE-A/Indonesia/5/2005 H5 TM+CT

MGTVNKPVVGVLMGFGIITGTLRITNPVRASVLRYYDDFHTDEDKLDNTNSVYEPYHSDHAESSWVNRGESS
 RKAYDHNSPYIWPRNDYDGFLENAHEHHGVYNQGRGIDSGERLMQPTQMSAQEDLGDDTGIHVIPITLNGDD
 RHKIVNVDQRQYGDVFKGDLNPKPQGQRLIEVSVEENHPFTLRAPIQRIYGVRYTETWSFLPSLTCTGDAA
 PAIQHICLKHTTCFQDVVDVDC AENTKEDQLAEISYRFQGGKEADQPWIVVNTSTLFDELELDPPEIEPG
 VLKVLRTKQYLGVIWNMRGSDGTSTYATFLVTWKGDEKTRNPTPAVTPQPRGAEEHWMNYHSHVFSVGD
 TFSLAMHLQYKIHEAPFDLLEWLYVPIDPTCQPMRLYSTCLYHPNAPQCLSHMNSGCTFTSPHLAQRVAS
 TVYQNCEHADNYTAYCLGISHMEPSFGLILHDGGTTLKFDVTPESLSGLYVFVYFNGHVEAVAYTVVSTV
 DHFVNAIEERGFPPTAGQPPATTKPKEITPVNPGTSPLLRQILSIYSTVASSLALAIMMAGLSLWMCSNGS
 LQCRICI*

Varicella Zoster virus (VZV)

Fig. 11A



Đối chứng dương (C+): VZL gE, Abcam, ab43050
 Đối chứng âm (C-): Dịch chiết từ lá được thâm nhiễm giả
 Kháng thể sơ cấp: protein kháng VZVgE mAb ở chuột, abcam, ab52549 (1 µg/ml)
 Kháng thể thứ cấp: Chuột kháng dê, JIR, 115-035-146 (1/10000)

Fig. 11B

Phương pháp loại trừ kích thước trên dịch chiết thô (946)

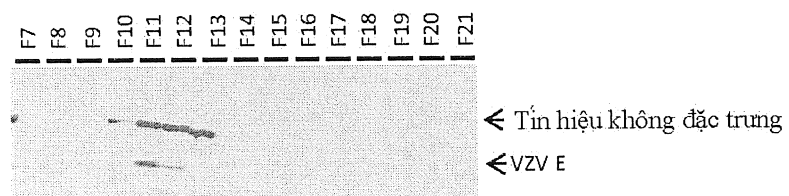


Fig. 12A

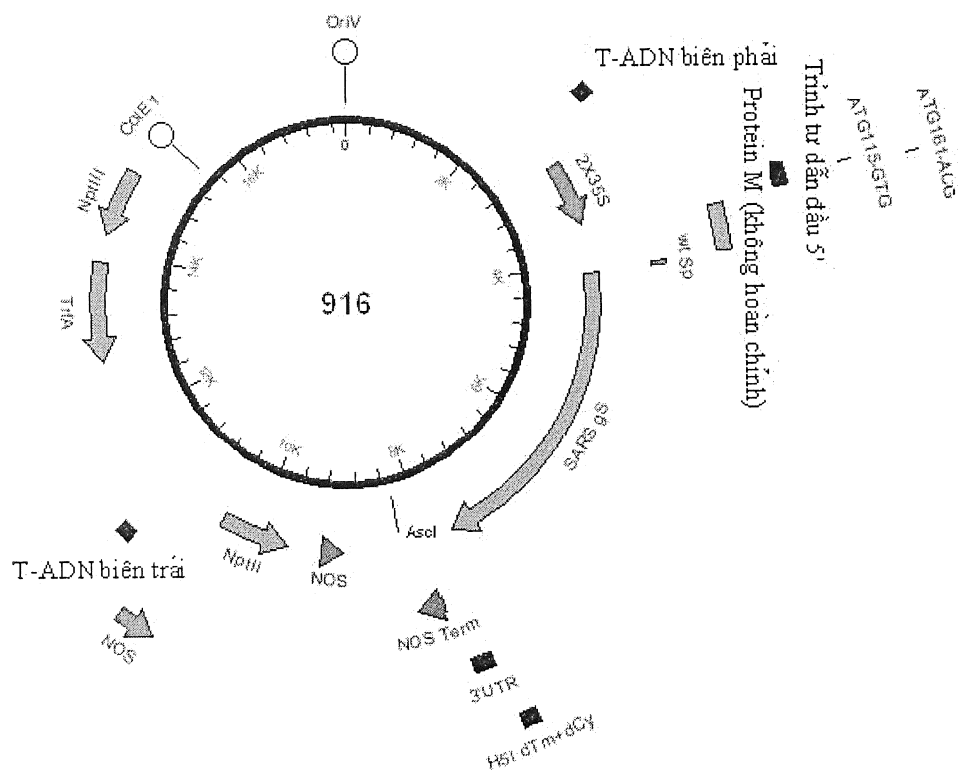


Fig. 12B, SEQ ID Số: 26

IF-wtSp-SARSgS.c

TGCCCAAATTTGTCGGGCCCATGTTTATTTTCTTATTATTCTTACTCTC

Fig. 12C, SEQ ID Số: 27

IF-H5dTm+SARSgS.r

AATTGACAGTATTTGAGGCCATTTAATATATTGCTCATATTTTCC

Fig. 12D, SEQ ID Số: 28

gen SARS gS được tổng hợp (tương ứng với nt 21492-25259 từ Số gia nhập ngân hàng gen AY278741.1)
(Peptit tín hiệu bám sinh được in đậm, vùng xuyên màng và vùng chắt nguyên sinh bám sinh được gạch dưới)

ATGTTTATTTCTTATTATTTCTTACTCTCACTAGTGGTAGTGACCTTGACCGGTGCACCACTTTTGATG
ATGTTCAAGCTCCTAATTACACTCAACATACTTCATCTATGAGGGGGGTTTACTATCCTGATGAAATTTT
AGATCAGACACTCTTATTTAACTCAGGATTTATTTCTTCCATTTTATTCTAATGTTACAGGGTTTCATACT
ATTAATCATACGTTTGGCAACCCTGTCATACCTTTTAAGGATGGTATTTATTTTGTGCCACAGAGAAAT
CAAATGTTGTCCGTGGTTGGGTTTTGGTTCTACCATGAACAACAAGTCACAGTCGGTGATTATTATTAA
CAATTCTACTAATGTTGTTATACGAGCATGTAACCTTTGAATTGTGTGACAACCCTTTCTTTGCTGTTTCTA
AACCCATGGGTACACAGACACATACTATGATATTCGATAATGCATTTAATTGCACTTTTCGAGTACATATC
TGATGCCTTTTCGCTTGATGTTTCAGAAAAGTCAGGTAATTTTAAACACTTACGAGAGTTTGTGTTTAAAA
ATAAAGATGGGTTTCTCTATGTTTATAAGGGCTATCAACCTATAGATGTAGTTCGTGATCTACCTTCTGG
TTTTAACACTTTGAAACCTATTTTTAAGTTGCCTCTTGGTATTAACATTACAAATTTTAGAGCCATTCTTAC
AGCCTTTTCACCTGCTCAAGACATTTGGGGCACGTCAGCTGCAGCCTATTTTGTGGCTATTTAAAGCC
AACTACATTTATGCTCAAGTATGATGAAAATGGTACAATCACAGATGCTGTTGATTGTTCTCAAAATCCA
CTTGCTGAACTCAAATGCTCTGTTAAGAGCTTTGAGATTGACAAAGGAATTTACCAGACCTCTAATTTCA
GGGTTGTTCCCTCAGGAGATGTTGTGAGATTCCCTAATATTACAACTTGTGTCCTTTTGGAGAGGTTTT
TAATGCTACTAAATCCCTTCTGTCTATGCATGGGAGAGAAAAAAATTTCTAATTGTGTTGCTGATTACT
CTGTGCTCTACAACCTCAACATTTTTTCAACCTTTAAGTGCTATGGCGTTTCTGCCACTAAGTTGAATGAT
CTTTGCTTCTCCAATGTCTATGCAGATTCTTTTGTAGTCAAGGGAGATGATGTAAGACAAATAGCGCCAG
GACAACTGGTGTTATTGCTGATTATAATTATAAATTGCCAGATGATTTTCATGGGTTGTGTCCTTGCTTG
GAATACTAGGAACATTGATGCTACTTCAACTGGTAATTATAATTATAAATATAGGTATCTTAGACATGGCA
AGCTTAGGCCCTTTGAGAGAGACATATCTAATGTGCCTTTCTCCCCTGATGGCAAACCTTGCACCCAC
CTGCTCTTAATTGTTATTGGCCATTAATGATTATGGTTTTTACACCACTACTGGCATTGGCTACCAACCT
TACAGAGTTGTAGTACTTTCTTTGAACTTTTAAATGCACCGGCCACGGTTTGTGGACCAAAATTATCCA
CTGACCTTATTAAGAACCAGTGTGTCAATTTTAAATTTAATGGACTCACTGGTACTGGTGTGTTAACTCCT
TCTTCAAAGAGATTTCAACCATTTCAACAATTTGGCCGTGATGTTTCTGATTTCACTGATTCCGTTTCGAG
ATCCTAAAACATCTGAAATATTAGACATTTCACTTGCTCTTTTGGGGGTGTAAGTGTAATTACACCTGG
AACAAATGCTTCATCTGAAGTTGCTGTTCTATATCAAGATGTTAACTGCACTGATGTTTCTACAGCAATTC
ATGCAGATCAACTCACACCAGCTTGGCGCATATATTCTACTGGAAACAATGATTCCAGACTCAAGCAG
GCTGTCTTATAGGAGCTGAGCATGTCGACACTTCTTATGAGTGCGACATTCCCTATTGGAGCTGGCATT
GTGCTAGTTACCATAACAGTTTCTTTATTACGTAGTACTAGCCAAAAATCTATTGTGGCTTATACTATGTCT
TTAGGTGCTGATAGTTCAATTGCTTACTCTAATAACACCATTGCTATACCTACTAACTTTTCAATTAGCAT
TACTACAGAAGTAATGCCTGTTTCTATGGCTAAAACCTCCGTAGATTGTAATATGTACATCTGCGGAGAT
TCTACTGAATGTGCTAATTTGCTTCTCCAATATGGTAGCTTTTGCACACAACCTAAATCGTGCCTCTCAG
GTATTGCTGCTGAACAGGATCGCAACACACGTGAAGTGTTGCTCAAGTCAAACAATGTACAAAACCC
CAACTTTGAAATATTTTGGTGGTTTTAATTTTCACAAATATTACCTGACCCTCTAAAGCCAACCTAAGAGG
TCTTTTATTGAGGACTTGCTCTTTAATAAGGTGACACTCGCTGATGCTGGCTTCATGAAGCAATATGGCG
AATGCCTAGGTGATATTAATGCTAGAGATCTCATTGTGCGCAGAAGTTCAATGGACTTACAGTGTTGC
CACCTCTGCTCACTGATGATATGATTGCTGCCTACACTGCTGCTCTAGTTAGTGGTACTGCCACTGCTG
GATGGACATTTGGTGCTGGCGCTGCTCTTCAAATACCTTTTGCTATGCAAATGGCATATAGGTTCAATG
GCATTGGAGTTACCCAAAAATGTTCTCTATGAGAACCAAAAAACAATCGCCAACCAATTTAACAAGGCGAT
TAGTCAAATTCAGAATCACTTACAACAACATCAACTGCATTGGGCAAGCTGCAAGACGTTGTTAACCA
GAATGCTCAAGCATTAAACACACTTGTTAAACAACCTAGCTCTAATTTTGGTGCAATTTCAAGTGTGCTA
AATGATATCCTTTGCGGACTTGATAAAGTCGAGGCGGAGGTACAAATTGACAGGTTAATTACAGGCAGA
CTTCAAAGCCTTCAAACCTATGTAACACAACAACCTAATCAGGGCTGCTGAAATCAGGGCTTCTGCTAAT
CTTGCTGCTACTAAAATGTCTGAGTGTGTTCTTGACAATCAAAAAGAGTTGACTTTTGTGGAAAGGGCT
ACCACCTTATGTCCTTCCCACAAGCAGCCCCGCATGGTGTGCTTCTACATGTCACGTATGTGCCAT

CCCAGGAGAGGAACTTCACCACAGCGCCAGCAATTTGTCATGAAGGCAAAGCATACTTCCCTCGTGAA
GGTGTGTTTTGTGTTAATGGCACTTCTTGGTTTATTACACAGAGGAACTTCTTTCTCCACAAATAATTAC
TACAGACAATACATTTGTCTCAGGAAATTGTGATGTCGTTATTGGCATCATTAAACAACACAGTTTATGAT
CCTCTGCAACCTGAGCTCGACTCATTCAAAGAAGAGCTGGACAAGTACTTCAAAAATCATACATCACCA
GATGTTGATCTTGGCGACATTTCAAGGCATTAACGCTTCTGTGTCACATTCAAAAAGAAATTGACCGC
CTCAATGAGGTCGCTAAAAATTTAAATGAATCACTCATTGACCTTCAAGAATTGGGAAAATATGAGCAAT
ATATTAAATGGCCTTGGTATGTTTGGCTCGGCTTCATTGCTGGACTAATTGCCATCGTCATGGTTACAAT
CTTGCTTTGTTGCATGACTAGTTGTTGCAGTTGCCTCAAGGGTGCATGCTCTTGTGGTTCTTGCTGCAA
GTTTGATGAGGATGACTCTGAGCCAGTTCTCAAGGGTGTCAAATTACATTACACATAA

Fig. 12E, SEQ ID Số: 29

SARSSgS+H5dTm.c

ATATATTAAATGGCCTCAAATACTGTCAATTTATTCAACAGTG

Fig. 12F, SEQ ID Số: 30

Caset biểu hiện số 916, từ PacI (ngược chiều với trình tự khởi động) đến Ascl (xuôi chiều ngay với trình tự kết thúc NOS). SARS gS-A/Indonesia/5/2005 H5 TM+CT được gạch dưới

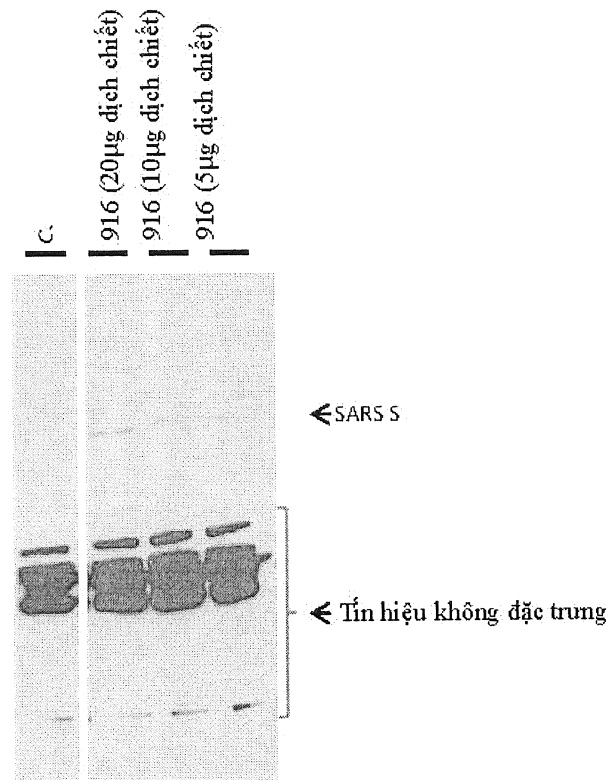
TTAATTAAGTCGACAAGCTTGCATGCCGGTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAA
TATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAA
CCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTC
CTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCA
AAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAA
GTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTC
TCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCAT
TGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCAT
TGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACC
CACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAT
CTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAG
TTCATTTTCATTTGGAGAGGTATTTAAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAACCCGAAC
CAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAGCAAACCTTCTCTTGTCTTTCTTGCGTGAGCGA
TCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTCACTGAAGCGAAATCAAAGAT
CTCTTTGTGGACACGTAGTGCAGGCCATTAAATAACGTGTACTTGTCTATTCTTGTGCGGTGTGGTCT
TGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTCAAGCCCATACATTACTTGTACGATTCTGCTGACTT
TCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTT
GCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTGAACAGAGTTTTCCCGTGGTTTTCGAACTTGGA
GAAAGATTGTTAAGCTTCTGTATTTCTGCCCAAATTTGTGGGCCCATGTTTATTTCTTATTATTTCTT
ACTCTCACTAGTGGTAGTGACCTTGACCGGTGCACCACTTTTGATGATGTTCAAGCTCCTAATTACACTC
AACATACTTCATCTATGAGGGGGTTTACTATCCTGATGAAATTTTTAGATCAGACACTCTTTATTTAACT
CAGGATTTATTTCTTCCATTTTATTCTAATGTTACAGGGTTTCATACTATTAATCATACGTTTGGCAACCC
TGTCATACCTTTTAAGGATGGTATTTATTTGCTGCCACAGAGAAATCAAATGTTGTCCGTGGTTGGGTT
TTTGGTTCTACCATGAACAACAAGTCACAGTCGGTGATTATTATAACAATTCTACTAATGTTGTTATACG
AGCATGTAACCTTTGAATTGTGTGACAACCTTTCTTTGCTGTTTCTAAACCCATGGGTACACAGACACAT
ACTATGATATTCGATAATGCATTTAATTGCACTTTCGAGTACATATCTGATGCCTTTTCGCTTGATGTTTC
AGAAAAGTCAGGTAATTTTAAACACTTACGAGAGTTTGTGTTTAAAAATAAAGATGGGTTTCTCTATGTTT

ATAAGGGCTATCAACCTATAGATGTAGTTCGTGATCTACCTTCTGGTTTTAACACTTTGAAACCTATTTTT
 AAGTTGCCTCTTGGTATTAACATTACAAATTTAGAGCCATTCTTACAGCCTTTTACCTGCTCAAGACAT
 TTGGGGCACGTGAGCTGCAGCCTATTTGTTGGCTATTTAAAGCCAACTACATTTATGCTCAAGTATGAT
 GAAAATGGTACAATCACAGATGCTGTTGATTGTTCTCAAATCCACTTGCTGAACTCAAATGCTCTGTTA
 AGAGCTTTGAGATTGACAAAGGAATTTACCAGACCTCTAATTTCAGGGTTGTTCCCTCAGGAGATGTTG
 TGAGATTCCCTAATATTACAACTTGTGTCCTTTGGAGAGGTTTTAATGCTACTAAATCCCTTCTGTC
 TATGCATGGGAGAGAAAAAAATTTCTAATTGTGTTGCTGATTACTCTGTGCTCTACAACCTCAACATTTTT
 TTCAACCTTTAAGTGCTATGGCGTTTCTGCCACTAAGTTGAATGATCTTTGCTTCTCCAATGTCTATGCA
 GATTCTTTTGTAGTCAAGGGAGATGATGTAAGACAAATAGCGCCAGGACAACTGGTGTATTGTCTGAT
 TATAATTATAAATTGCCAGATGATTTTCATGGGTTGTGTCCTTGCTTGGAATACTAGGAACACTGATGCTA
 CTTCAACTGGTAATTATAATTATAAATATAGGTATCTTAGACATGGCAAGCTTAGGCCCTTTGAGAGAGA
 CATATCTAATGTGCCTTTCTCCCCTGATGGCAAACCTTGACCCCACTGCTCTTAATTGTTATTGGCCA
 TTAATGATTATGGTTTTTACACCACTACTGGCATTGGCTACCAACCTTACAGAGTTGTAGTACTTTCTTT
 TGAACTTTTAATGCACCGGCCACGGTTTTGTGGACCAAAATATCCACTGACCTTATTAAGAACCAGTGT
 GTCATTTTTAATTTAATGGACTCACTGGTACTGGTGTGTTAACTCCTTCTCAAAGAGATTTCAACCATT
 TCAACAATTTGGCGTGATGTTTCTGATTTCACTGATTCCGTTTCGAGATCCTAAAACATCTGAAATATTA
 GACATTTACCTTGCTCTTTTGGGGGTGTAAGTGTAATTACACCTGGAACAAATGCTTCATCTGAAGTTG
 CTGTTCTATATCAAGATGTTAACTGCACTGATGTTTCTACAGCAATTCATGCAGATCAACTCACACCAGC
 TTGGCGCATATATTCTACTGGAAACAATGTATTCCAGACTCAAGCAGGCTGTCTTATAGGAGCTGAGCA
 TGTGACACTTCTTATGAGTGCGACATTCTATTGGAGCTGGCATTGTGCTAGTTACCATACAGTTTCT
 TTATTACGTAGTACTAGCCAAAAATCTATTGTGGCTTATACTATGCTTTAGGTGCTGATAGTTCAATTGC
 TTAATCTAATAACACCATTGCTATACCTACTAACTTTTCAATTAGCATTACTACAGAAGTAATGCCTGTTT
 CTATGGCTAAAACCTCCGTAGATTGTAAATGTACATCTGCGGAGATTCTACTGAATGTGCTAATTTGCT
 TCTCCAATATGGTAGCTTTTGCACACAACCTAAATCGTGCCTCTCAGGTATTGCTGCTGAACAGGATCG
 CAACACACGTGAAGTGTTGCTCAAGTCAAACAAATGTACAAAACCCCACTTTGAAATATTTTGGTGGT
 TTTAATTTTACAAAATATTACCTGACCTCTAAAGCCAACTAAGAGGTCTTTTATTGAGGACTTGCTCTT
 TAATAAGGTGACACTCGCTGATGCTGGCTTCATGAAGCAATATGGCGAATGCCTAGGTGATTAATGC
 TAGAGATCTCATTGTGCGCAGAAGTTCAATGGACTTACAGTGTGCGCACCTCTGCTCACTGATGATAT
 GATTGCTGCTACACTGCTGCTCTAGTTAGTGCTACTGCCACTGCTGGATGGACATTTGGTGTGGCG
 CTGCTCTTCAAATACCTTTTGTCTATGCAAATGGCATATAGGTTCAATGGCATTGGAGTTACCCAAAATGT
 TCTCTATGAGAACCAAAAAACAAATCGCCAACCAATTTAACAAGGCGATTAGTCAAATCAAGAATCACTT
 ACAACAACATCAACTGCATTGGGCAAGCTGCAAGACGTTGTTAACCAGAATGCTCAAGCATTAAACACA
 CTTGTTAAACAACCTTAGCTCTAATTTTGGTGCAATTTCAAGTGTGCTAAATGATATCCTTTGCGGACTTGA
 TAAAGTCGAGGCGGAGGTACAAATTGACAGGTTAATTACAGGCAGACTTCAAAGCCTTCAAACCTATGT
 AACACAACAACCTAATCAGGGCTGCTGAAATCAGGGCTTCTGCTAATCTTGCTGCTACTAAAATGTCTGA
 GTGTGTTCTTGACAATCAAAAAGAGTTGACTTTTGTGGAAAGGGCTACCACCTTATGTCCTTCCCACA
 AGCAGCCCCGCATGGTGTTGTCTTCTACATGTCACGTATGTGCCATCCCAGGAGAGGAACCTCACCA
 CAGCGCCAGCAATTTGTCATGAAGGCAAAGCATACTTCCCTCGTGAAGGTGTTTTTGTGTTAATGGCA
 CTTCTTGTTTATTACACAGAGGAACCTTCTTTTCTCCACAAATAATTACTACAGACAATACATTTGTCTCA
 GGAAATTGTGATGTCGTTATTGGCATCATTAACAACACAGTTTATGATCCTCTGCAACCTGAGCTCGACT
 CATTCAAAGAAGAGCTGGACAAGTACTTCAAAAATCATACATCACCAGATGTTGATCTTGCGGACATTTT
 AGGCATTAACGCTTCTGTCGTCAACATTCAAAAAGAAATTGACCGCCTCAATGAGGTGCTAAAAATTTA
 AATGAATCACTCATTGACCTTCAAGAATTGGGAAAATATGAGCAATATATTAAATGGCCTCAAATACTGT
 CAATTTATTCAACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGT
 GCTCCAATGGATCGTTACAATGCAGAAATTTGCATTTAAAGGCCTATTTTCTTTAGTTTGAATTTACTGTTA
 TTCGGTGTGCATTTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAAT
 TTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATTTATTA
 AAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAA
 TAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGT
 TAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCG
 CAATTATACATTTAATACGCGATAGAAAACAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCG
 GTGTCATCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCC

Fig. 12G, SEQ ID Số: 31

Trình tự axit amin của SARS gS- A/Indonesia/5/2005 H5 TM+CT

MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTLQDLFL
 PFYSNVTGFHTINHTFGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNS
 TNVIRACNFELCDNPFFAVSKPMGTQTHTMIFDNAFNCTFEYISDAFSLDVSEKSGNF
 KHLREFVFKNKDGLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILTAFS
 PAQDIWGTSAAYFVGYLKPTTFMLKYDENGITITDAVDCSQNPLAELKCSVKSFEIDKGIY
 QTSNFRVVPSPGDVVRFPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNST
 FFSTFKCYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFM
 GCVLAWNTRNIDATSTGNYNKYRYLRHGKLRPFERDISNVPFSPDGKPCPTPALNCY
 WPLNDYGFYTTTGIGYQPYRVVLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTG
 TGVLTTPSSKRFQPFQFGRDVSDFTDSVRDPKTSEILDSPCSFGGVSVITPGTNASSE
 VAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDIPI
 GAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMPVSMA
 KTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYK
 TPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLI
 CAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGI
 GVTQNVLYENQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSS
 NFGAISSVLNDILSRDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKM
 SECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKA
 YFPREGVVFNGTSWFITQRNFFSPQIITDNTFVSGNCDVVIGIINNTVYDPLQPELDSF
 KEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQ
 YIKWPQILSIYSTVASSLALAIMMAGLSLWMCSNGSLQCRICI*



Đối chứng âm (C-): Dịch chiết từ lá được thâm nhiễm giả
 Kháng thể sơ cấp: kháng thể đa dòng IgG1 ở thỏ, Imgenex, IMG-690 (2µg/ml)
 Kháng thể thứ cấp: IgG thỏ kháng chuột, JIR, 111-035-144 (1/10000)

Fig. 12H

Fig. 13A (SEQ ID Số: 43)

IF-Opt_EboGP.s2+4c

TCTCAGATCTTCGCCATCCCTCTGGGAGTGATCCACAATTCGACT

Fig. 13B (SEQ ID Số: 44)

H5iTMCT+Opt_EboGP.r

TTGAATAAATTGACAGTATCTGGCGCCATCCGGTCCACCAGTTATCGT

Fig. 13C (SEQ ID Số: 45)

Gen GP tổng hợp được tối ưu hóa (tương ứng với nt 6039-8069 từ Số gia nhập ngân hàng gen AY354458 với trình tự gen loại dại. Gen được tối ưu hóa để sử dụng codon và hàm lượng GC, sự loại bỏ các vị trí cắt bí ẩn, trình tự Shine-Delgarno, trình tự khởi ổn định ARN và vị trí đi vào ribosom prokaryot) (Peptit tín hiệu bám sinh được in đậm, vùng xuyên màng bám sinh và vùng chất nguyên sinh được gạch dưới)

ATGGGGGTGACCGGAATCCTTCAGCTCCCTCGCGATAGGTTCAAACGGACGTCCTTCTTTCTCTGG
GTGATCATCCTGTTTCAGCGCACCTTCTCAATCCCTCTGGGAGTGATCCACAATTCGACTCTCCAGGTG
 AGCGAGGTCGACAAACTCGTGTGCCGCGATAAGTTGTCCAGCACCAATCAGCTCAGAAGTGTTGGCCTCA
 ACCTGGAGGGGAATGGGGTCGCCACAGACGTGCCATCCGCGACCAAGCGCTGGGGCTTTCGGTCGGGTG
 TTCCCCCTAAAGTCGTGAATTACGAGGCCGGGGAGTGGGCTGAAAAGTCTACAACCTGGAAATCAAGA
 AACCCGACGGCAGCGAGTGTCTGCCCGCGGCCCCAGACGGGATCAGGGGTTTCCCGAGGTGCCGGTACG
 TGCACAAGGTCTCGGGAACCGGCCCGTGCGCCGGCGATTTTGCCTTCCATAAGGAGGGCGCATTCTTTCT
 CTACGATAGACTGGCCTCCACGGTCATCTATCGCGGCACCACTTCGCGGAGGGGGTGGTGGCATTCTCTC
 ATCCTGCCGACGGCTAAGAAAGATTTCTTCAGCTCCCACCCTCTCAGGGAACCAAGTCAACGCCACAGAGG
 ATCCCTCCTCGGGATATTACAGCACCACAATTCGGTACCAGGCCACAGGCTTCGGGACCAATGAGACTGA
 GTACCTGTTTCGAGGTGGACAACCTAACATACGTGCAGCTCGAATCGCGGTTTACCCCGCAGTTCTTTTGC
 AACTTAACGAGACGATCTATACAGTGGTAAACGGAGCAACACGACCGGGAAGCTCATCTGGAAAGTAA
 ACCCGGAGATCGACACCACGATTGGAGAGTGGGCTTTCTGGGAAACCAAGAAAAACCTGACCCGGAAGA
 TCAGGTCGGAGGAGTTGAGCTTCACTGCCGTCTCCAATAGAGCAAAAAACATCAGCGGCCAGAGCCCTG
 CCCGACTTCCAGCGACCCAGGTACCAACACCACGACAGAGGACCACAAGATTATGGCCTCGGAGAATT
 CCTCTGCAATGGTCCAGGTGCATAGCCAGGGGCGCGAAGCTGCCGTCTCGCACCTGACAACCTTGGCAAC
 GATTTCCACCAGCCCACAACCACCCACGACAAAGCCCGGGCCCGACAACCTCCACCCATAATACCCCGGTC
 TATAAGCTGGACATTTCCGAAGCTACGCAGGTGGAGCAGCACCACCGGAGGACCGACAACGACTCAACA
 GCTAGCGATACCCCTCCCGCCACCACGGCAGCGGGGCCACCCAAGGCTGAGAACACCAATACGAGCAAG
 GGCACCGACCTCCTGGACCCCGCGACTACCACGAGCCCCCAGAATCACAGCGAAACGGCGGGCAACAAC
 AACACACACCATCAGGATACTGGGGAGGAGTGGGCCTCCAGCGGAAAACCTGGGCCTGATCACAACACCC
 ATAGCCGGGGTCGCCGGGCTGATCACTGGAGGCCGGAGGGCACGCAGAGAGGCCATTGTGAACGCCAG
 CCCAAATGCAACCCAAACCTCCACTACTGGACCACGCAGGACGAGGGGGCCGCTATCGGCCTGGCCTGG
 ATCCCATACTTTGGGCCCCGCGCTGAGGGCATATACACGGAGGGCCTCATGCACAATCAGGACGGACTG
 ATCTGCGGACTCCGCCAGCTTGCCAACGAGACCACTCAGGCGTTGCAGCTGTTTCTGCGGGCCACTACCG
 AGCTCAGGACGTTTCAGCATTCTGAATCGGAAGGCAATCGATTTCTACTCCAGCGGTGGGGCGGGACATG
 CCACATCCTGGGACCCGATTGTTGCATCGAGCCCCACGACTGGACCAAGAACATTACAGATAAAATCGA
 CCAGATCATTCATGATTTCTGTGGACAAGACACTGCCGGACCAGGGGGACAACGATAACTGGTGGACCGG

V83777WO

ATGGCGCCAGTGGATCCCAGCCGGGATTGGCGTGACAGGTGTCATTATCGCCGTGATCGCCCTGTTTTC
 ATTTGCAAATTCGTTTTCTAG

Fig 13D (SEQ ID Số: 46)

Opt_EboGP+H5iTMCT.c

GGATGGCGCCAGATACTGTCAATTTATTCAACAGTGGCGAGTTCC

Fig. 13E

Sự thể hiện dưới dạng sơ đồ của sản phẩm thiết kế 1192. Vị trí enzym giới hạn SacII và StuI được sử dụng cho sự tuyến tính hóa plasmit được chú thích trên đồ biểu diễn.

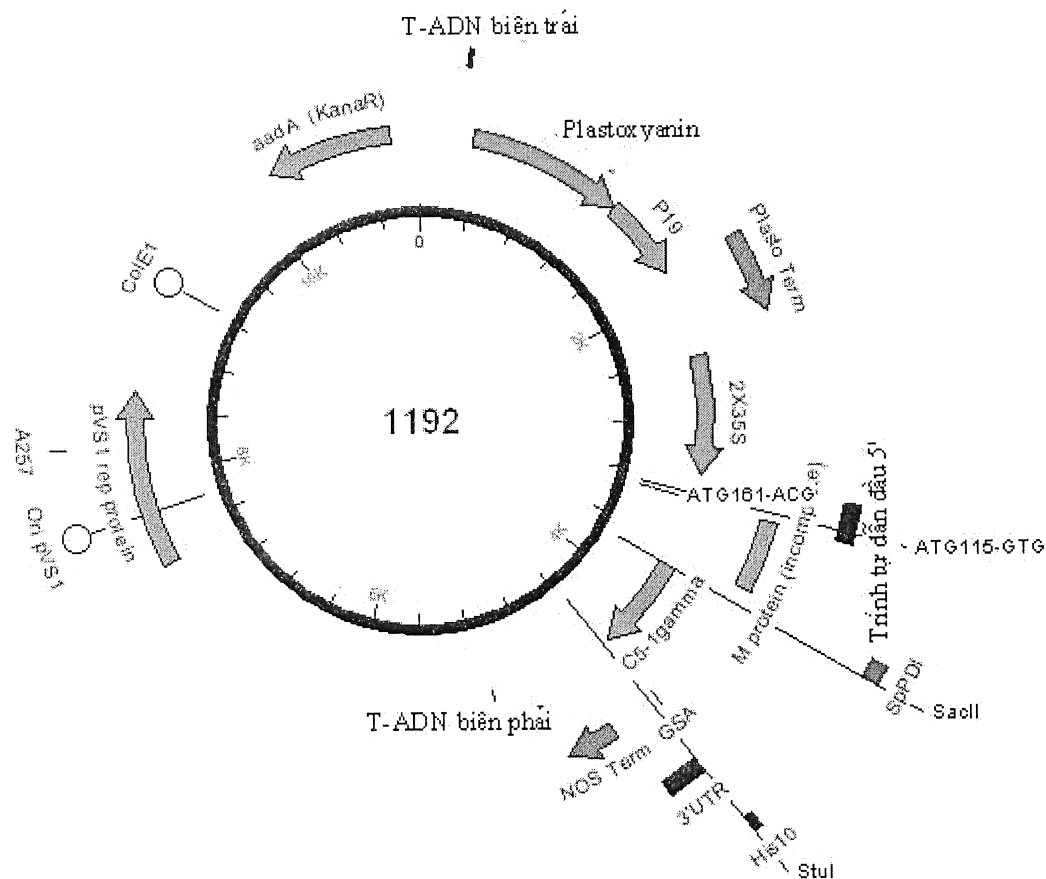


Fig. 13F, SEQ ID Số: 47

Sản phẩm thiết kế 1192 từ đường biên t-ADN từ trái sang phải (được gạch dưới). 2X35S/CPMV-HT/
PDISP/NOS với caset biểu hiện chất ức chế bất hoạt plastoxyanin-P19-Plastoxyanin

TGGCAGGATATATTGTGGTGTAAACAAATTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTT
 ATGTACTGAATTAACGCCGAATCCCGGGCTGGTATATTTATATGTTGTCAAATAACTCAAAAACCATAAA
 AGTTTAAAGTTAGCAAGTGTGTACATTTTACTTGAACAAAAATATTCACCTACTACTGTTATAAATCATT
 TTAACATTAGAGTAAAGAAATATGGATGATAAGAACAAGAGTAGTGATATTTTGACAACAATTTTGTG
 CAACATTTGAGAAAATTTTGTGTTCTCTCTTTTCATTGGTCAAAAACAATAGAGAGAGAAAAAGGAAGA
 GGGAGAATAAAAACATAATGTGAGTATGAGAGAGAAAGTTGTACAAAAGTTGTACCAAAATAGTTGTAC
 AAATATCATTGAGGAATTTGACAAAAGCTACACAAATAAGGGTTAATTGCTGTAAATAAATAAGGATGA
 CGCATTAGAGAGATGTACCATTAGAGAATTTTGGCAAGTCATTAAGAAAGAAAGAATAAATTATTTTAA
 AATTAAGAGTTGAGTCATTTGATTAAACATGTGATTATTTAATGAATTGATGAAAGAGTTGGATTAAAGT
 TGTATTAGTAATTAGAATTTGGTGTCAAATTTAATTTGACATTTGATCTTTTCCTATATATTGCCCATAGA
 GTCAGTTAACTCATTTTTATATTTTCATAGATCAAATAAGAGAAATAACGGTATATTAATCCCTCCAAAA
 AAAAAACGGTATATTTACTAAAAAATCTAAGCCACGTAGGAGGATAACAGGATCCCCGTAGGAGGATA
 ACATCCAATCCAACCAATCACACAATCCTGATGAGATAACCCACTTTAAGCCACGCATCTGTGGCACA
 TCTACATTATCTAAATCACACATTCTTCCACACATCTGAGCCACACAAAAACCAATCCACATCTTTATCAC
 CCATTCTATAAAAAATCACACTTTGTGAGTCTACACTTTGATTCCCTTCAAACACATACAAAGAGAAGAG
 ACTAATTAATTAATTAATCATCTTGAGAGAAAATGGAACGAGCTATACAAGGAAACGACGCTAGGGAAC
 AAGCTAACAGTGAACGTTGGGATGGAGGATCAGGAGGTACCACTTCTCCCTTCAAACCTCCTGACGAAAG
 TCCGAGTTGGACTGAGTGGCGGTACATAACGATGAGACGAATTCGAATCAAGATAATCCCTTGGTTTC
 AAGGAAAGCTGGGGTTTCGGGAAAAGTTGTATTTAAGAGATATCTCAGATACGACAGGACGGAAGCTTCA
 CTGCACAGAGTCCTTGGATCTTGGACGGGAGATTCCGTTAACTATGCAGCATCTCGATTTTTCGGTTTCGA
 CCAGATCGGATGTACCTATAGTATTCGGTTTCGAGGAGTTAGTATCACCGTTTCTGGAGGGTCGCGAACT
 CTTCAGCATCTCTGTGAGATGGCAATTCGGTCTAAGCAAGAACTGCTACAGCTTGCCCCAATCGAAGTGG
 AAAGTAATGTATCAAGAGGATGCCCTGAAGGTACTCAAACCTTCGAAAAAGAAAGCGAGTAAGTTAAAA
 TGCTTCTTCGTCTCCTATTTATAATATGGTTTGTATTGTTAATTTTGTCTTGTAGAAGAGCTTAATTAAT
 CGTTGTTGTTATGAAATACTATTTGTATGAGATGAACTGGTGTAAATGTAATTCATTTACATAAGTGGAGTC
 AGAATCAGAATGTTTCCTCCATAACTAACTAGACATGAAGACCTGCCGCTACAATTGTCTTATATTTGA
 ACAACTAAAATTGAACATCTTTTGCCACAACCTTTATAAGTGGTTAATATAGCTCAAATATATGGTCAAGTT
 CAATAGATTAATAATGGAAATATCAGTTATCGAAATTCATTAACAATCAACTTAACGTTATTAACACTA
 ATTTTATATCATCCCTTTGATAAATGATAGTACACCAATTAGGAAGGAGCATGCTCGCCTAGGAGATTG
 TCGTTTCCCGCCTTCAGTTTGCAAGCTGCTCTAGCCGTGTAGCCAATACGCAAACCGCCTCTCCCCGCGCG
 TTGGGAATTACTAGCGCGTGTGACAAGCTTGCATGCCGGTCAACATGGTGGAGCACGACACACTTGTCT
 ACTCAAAAATATCAAAGATACAGTCTCAGAAAGACCAAGGGCAATTGAGACTTTTCAACAAAGGGTAA
 TATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGA
 AGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGT
 GGTCCTCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCA
 AAGCAAGTGGATTGATGTGATAACATGGTGGAGCACGACACACTTGTCTACTCAAAAATATCAAAGAT
 ACAGTCTCAGAAGACCAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGA
 TTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCC
 ATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCC
 CACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTG
 ATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCTTCTCTATATAAGG
 AAGTTCATTTTATTTGGAGAGGTATTAATCTTAATAGGTTTTGATAAAAGCGAACGTGGGGAAACCCG
 AACCAACCTTCTTCTAACTCTCTCTCATCTCTCTTAAAGCAAACCTCTCTTGTCTTTCTTGCCTGAGC
 GATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACAACGTTTTCTTTCACTGAAGCGAAATCAAAG
 ATCTCTTTGTGGACACGTAGTGGCGGCCATTAAATAACGTGTACTTGTCTTATTCTTGTCTGCTGCTG
 TGGGAAAAGAAAGCTTGCTGGAGGCTGCTGTTGAGCCCCATACATTACTTGTACGATTCTGCTGACTTTC
 GGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTGTTGCCTGTACTTCTTTCTTCTTCTTCTGCTG
 ATTGTTCTATAAGAAATCTAGTATTTTCTTGAAACAGAGTTTTCCCGTGGTTTTCGAACTTGGAGAAAG
 ATTGTTAAGCTTCTGTATATTCTGCCAAATTTGTGCGGCCCATGGCGAAAAACGTTGCGATTTTCGGCTT

ATTGTTTCTCTTCTTGTGTGGTTCTTCTCAGATCTTCGCCGCGGCTCCTCAGCCAAAACGACACCCCCA
TCTGTCTATCCACTGGCCCCTGGATCTGCTGCCAAACTAACTCCATGGTGACCCTGGGATGCCTGGTCAA
GGGCTATTTCCCTGAGCCAGTGACAGTGACCTGGAACCTCTGGATCCCTGTCCAGCGGTGTGCACACCTTC
CCAGCTGTCTGCACTGTGACCTCTACACTCTGAGCAGCTCAGTGACTGTCCCTCCAGCACCTGGCCCA
GCGAGACCGTCACCTGCAACGTTGCCACCCGCGCCAGCAGACCAAGGTGGACAAGAAAATTGTGCCA
GGGATTGTGGTTGTAAGCCTTGCATATGTACAGTCCCAGAAGTATCATCTGTCTTCATCTTCCCCCAAAG
CCCAAGGATGTGCTCACCATTACTCTGACTCCTAAGGTCACGTGTGTTGTGGTAGACATCAGCAAGGATG
ATCCCAGAGTCCAGTTCACTGTGTTGTAGATGATGTGGAGGTGCACACAGCTCAGACGCAACCCCGGGA
GGAGCAGTTCAACAGCACTTTCCGCTCAGTCAGTGAACCTCCCATCATGCACCAGGACTGGCTCAATGGC
AAGGAGCGATCGCTCACCATCACCATCACCATCACCATCACCATTAAAGGCCTATTTTCTTTAGTTTGAAT
TTACTGTTATTTCGGTGTGCATTTCTATGTTTGGTGAGCGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGT
AATTTAATTTCTTTGTGAGCTCCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAAATT
TATTAATAAAAAAAAAAAAAAAAAAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAAC
ATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCATGATTATCATATAATTTCTGTT
GAATTACGTTAAGCATGTAATAATTAAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTA
GAGTCCCGCAATTATACATTTAATACGCGATAGAAAAACAAATATAGCGCGCAAACTAGGATAAATTAT
CGCGCGCGGTGTCTATGTTACTAGATCTCTAGAGTCTCAAGCTTGGCGCGCCACGTGACTAGTGGC
ACTGGCCGTGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCAACCTTAATCGCCTTGCAGCA
CATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCCGACCGATCGCCCTTCCCAACAGTTGCGCA
GCCTGAATGGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGATTGTCGTTTTCCCGCCTTCAGTTTAACT
ATCAGTGTGTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTTTA

Fig. 13G, SEQ ID Số: 48

Caset biểu hiện số 1366 từ trình tự khởi động 2X35S đến trình tự kết thúc NOS. PDISP-GP từ Zaire 95 Ebolavirut-H5 TM+CT từ trình tự A/Indonesia/5/2005 được gạch dưới.

GTCTAACATGGTGGAGCAGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAA
AGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCT
GTCACTTTATTGTGAAGATAGTGGAAGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAA
AGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAGATGGACCCCCACCCACGAGGAGCATCG
TGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG
ACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTGAGACTTTTCA
ACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATA
GTGGAAGAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCC
TCTGCCGACAGTGGTCCCAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCA
ACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACT
ATCCTTCGCAAGACCCTTCTCTATATAAGGAAGTTCATTTCAATTTGGAGAGGTATTAAAATCTTAATAGG
TTTTGATAAAAGCGAACGTGGGGAAACCCGAACCAAACCTTCTTCTAAACTCTCTCTCATCTCTCTTAAAG
CAAACCTCTCTCTTGTCTTTCTTGCGTGAGCGATCTTCAACGTTGTCAGATCGTGCTTCGGCACCAGTACA
ACGTTTTCTTCACTGAAGCGAAATCAAAGATCTCTTTGTGGACACGTAGTGCGGCGCCATTAAATAACG
TGTACTTGTCTATTTCTTGTCTGGTCTTGGGAAAAAGAAAGCTTGTGGAGGCTGCTGTTTACGCCCCA
TACATTACTTGTACGATTCTGCTGACTTTCGGCGGGTGCAATATCTCTACTTCTGCTTGACGAGGTATTG
TTGCCGTGACTTCTTTCTTCTTCTTCTTCTGCTGATTGGTTCTATAAGAAATCTAGTATTTTCTTTGAAACAGA
GTTTTCCCGTGTTTTCGAACTTGGAGAAAGATTGTTAAGCTTCTGTATATTCTGCCCAAATTTGTCTGGGC
CCATGGCGAAAAACGTTGCGATTTTCGGCTTATTGTTTTCTCTTGTGTTGTTCTCTTCTCAGATCTTCG
CCATCCCTCTGGGAGTGATCCACAATTGCACTCTCCAGGTGAGCGAGTGCACAAACTCGTGTGCCCGGA
TAAGTTGTCCAGCACCAATCAGCTCAGAAGTTTGGCCTCAACCTGGAGGGGAATTGGGGTCGCCACAGA
CGTGCCATCCGCGACCAAGCGCTGGGGCTTTGGTGGGTTTTCCTTAAAGTCGTGAATTACGAGGCC
GGGGAGTGGGCTGAAAACCTGCTACAACTGGAAATCAAGAAACCCGACGGCAGCGAGTGTCTGCCCCGCG
GCCCCAGACGGGATCAGGGGTTTCCCGAGGTGCCGGTACGTGCACAAGGTCTCGGGAACCGGCCCGTGC
GCCGGCGATTTTGCCTTCCATAAGGAGGGCGCATTCTTTCTCTACGATAGACTGGCCTCCACGGTCATCTA

TCGCGGCACCACTTCGCGGAGGGGGTGGTGGCATTCTCATCCTGCCGCAGGCTAAGAAAGATTCTTC
 AGCTCCCACCCTCTCAGGGAACCACTCAACGCCACAGAGGATCCCTCCTCGGGATATTACAGCACCACAA
 TTCGGTACCAGGCCACAGGCTTCGGGACCAATGAGACTGAGTACCTGTTTCGAGGTGGACAACCTAACATA
 CGTGCAGCTCGAATCGCGGTTTACCCCGCAGTTCCCTTTTGCAACTTAACGAGACGATCTATACCAGTGGT
 AAACGGAGCAACACGACCGGGAGCTCATCTGGAAAGTAAACCCGGAGATCGACACCACGATTGGAGA
 GTGGGCTTTCTGGGAAACCAAGAAAAACCTGACCCGGAAGATCAGGTCCGAGGAGTTGAGCTTCACTGC
 CGTCTCCAATAGAGCAAAAAACATCAGCGGCCAGAGCCCTGCCCGGACTTCCAGCGACCCAGGTACCAA
 CACCACGACAGAGGACCACAAGATTATGGCCTCGGAGAATTCTCTGCAATGGTCCAGGTGCATAGCCA
 GGGGCGCGAAGCTGCCGTCTCGCACCTGACAACCTTGGAACGATTTCCACCAGCCCACAACCACCCACG
 ACAAAGCCCCGGGCCCCGACAACCTCCACCCATAATACCCCGGTCTATAAGCTGGACATTTCCGAAGCTACGC
 AGGTGGAGCAGCACCACCGGAGGACCGACAACGACTCAACAGCTAGCGATACCCCTCCCGCCACCACGG
 CAGCGGGCCCCACCAAGGCTGAGAACACCAATACGAGCAAGGGCACCGACCTCCTGGACCCCGCGACTA
 CCACGAGCCCCCAGAATCACAGCGAAACGGCGGGCAACAACAACACACACCATCAGGATACTGGGGAG
 GAGTCGGCCTCCAGCGGAAAACCTGGGCCTGATCACAACACCATAGCCGGGGTTCGCCGGGCTGATCACT
 GGAGGCCGGAGGGCACGCAGAGAGGCCATTGTGAACGCCAGCCCAATGCAACCCAAACCTCCACTAC
 TGGACCACGCAGGACGAGGGGGCCGCTATCGGCCTGGCCTGGATCCCATACTTTGGGCCCCGCGCTGAG
 GGCATATACACGGAGGGCCTCATGCACAATCAGGACGGACTGATCTGCGGACTCCGCCAGCTTGCCAAC
 GAGACCACTCAGGCGTTGCAGCTGTTTCTGCGGGCCACTACCGAGCTCAGGACGTTTCAAGCATTCTGAATC
 GGAAGGCAATCGATTTCCCTACTCCAGCGGTGGGGCGGGACATGCCACATCCTGGGACCCGATTGTTGCAT
 CGAGCCCCACGACTGGACCAAGAACATTACAGATAAAATCGACCAGATCATTATGATTTCTGGACAA
 GACACTGCCGGACCGGGGGACAACGATAAAGTGGTGGACCGGATGGCGCCAGATACTGTCAATTTATTTC
 AACAGTGGCGAGTTCCCTAGCACTGGCAATCATGATGGCTGGTCTATCTTTATGGATGTGCTCCAATGGA
 TCGTTACAATGCAGAAATTTGCATTTAAAGGCCTATTTCTTTAGTTTGAATTTACTGTTATTTCGGTGTGCAT
 TTCTATGTTTGGTGAGCGGTTTTCTGTGCTCAGAGTGTGTTTATTTTATGTAATTTAATTTCTTTGTGAGCT
 CCTGTTTAGCAGGTCGTCCCTTCAGCAAGGACACAAAAAGATTTTAATTTTATTAATAAAAAAAAAAAAAA
 AAGACCGGAATTCGATATCAAGCTTATCGACCTGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTA
 AGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAAATTTCTGTTGAATTACGTTAAGCATGTAAT
 AATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTATGATTAGAGTCCCGCAATTATACATTT
 AATACGCGATAGAAAACAAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCGGTGTCATCTATGT
 TACTAGAT

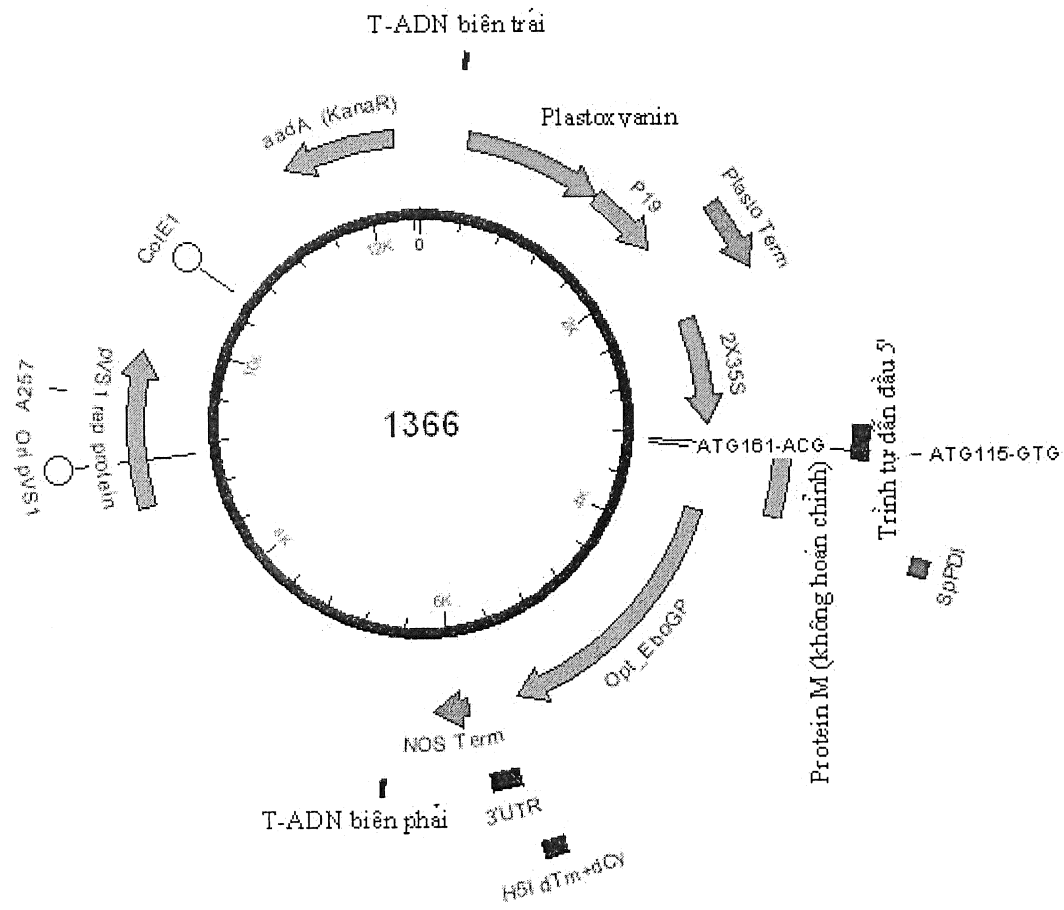
Fig. 13H, SEQ ID Sô: 49

Trình tự axit amin của PDISP-GP từ Zaire 95 Ebolavirut-H5 TM+CT từ A/Indonesia/5/2005

MAKNVAIFGLLFSLVLVPSQIFAIPLGVIHNSTLQVSEVDKLVCRDKLSSTNQLRVGLN
 LEGNGVATDVPSATKRWGFSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPD
 GIRGFPRCRYVHKVSGTGPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILP
 QAKKDFFSHPLREPVNATEDPSSGYYSTTIRYQATGFGTNETEYLFEVDNLTYVQLESR
 FTPQFLLQLNETIYTSGKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTKIRSEELSFT
 AVSNRAKNISGQSPARTSSDPGTNTTTEDHKIMASENSSAMVQVHSQGREAAVSHLTTL
 ATISTSPQPPTTKPGPDNSTHNTVPYKLDISEATQVEQHRRTDNDSTASDTPPATTAAGP
 PKAENTNTSKGTDLLDPATTTSPQNHSETAGNNNTHHQDTGEESASSGKLGLITNTIAGV
 AGLITGRRARREAIVNAQPKCNPNLHYWTTQDEGAAGLAWIPYFGPAEAGIYTEGLM
 HNQDGLICGLRQLANETTQALQLFLRATTELRTFSILNRKAIDFLLQRWGGTCHILGPDC
 CIEPHDWTKNITDKIDQIHDFVDKTLPDQGDNDNWWTGWQRILSIYSTVASSLALAIMM
 AGLSLWMCSNGSLQCRICI*

Fig. 13I

Sự thể hiện dưới dạng sơ đồ của sản phẩm thiết kế số 1366



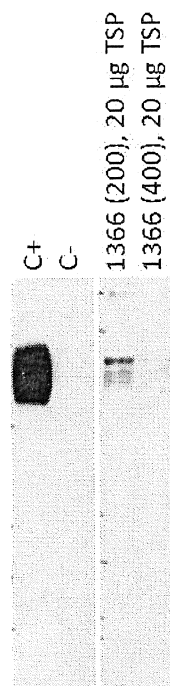


Fig. 13J

DANH MỤC TRÌNH TỰ

<110> Medicago Inc.
D'Aoust, Marc-Andre
Couture, Manon
Lavoie, Pierre-Olivier
Vezina, Louis-Philippe

<120> PHƯƠNG PHÁP SẢN XUẤT HẠT TƯƠNG TỰ VIRUT (VLP) Ở THỰC VẬT VÀ HẠT TƯƠNG TỰ VIRUT ĐƯỢC SẢN XUẤT BẰNG PHƯƠNG PHÁP NÀY

<130> V83777WO

<150> US 61/426,401
<151> 2010-12-22

<150> US 61/496,371
<151> 2011-06-13

<160> 49

<170> PatentIn version 3.5

<210> 1
<211> 2376
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Trình tự axit nucleic liên ứng của HIV ConS (denta)CFI

<400> 1
atgcgcgtgc gcggcatcca gcgcaactgc cagcacctgt ggcgctgggg caccctgatc 60
ctgggcatgc tgatgatctg ctccgccgcc gagaacctgt gggtgaccgt gtactacggc 120
gtgcccgtgt ggaaggaggc caacaccacc ctgttctgcg cctccgacgc caaggcctac 180
gacaccgagg tgcacaacgt gtggggccacc cagcctgcg tgcccaccga cccaacccc 240
caggagatcg tgctggagaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg 300
gagcagatgc acgaggacat catctccctg tgggaccagt ccctgaagcc ctgctgaag 360
ctgaccccc tgtgctgac cctgaactgc accaacgtga acgtgaccaa caccaccaac 420
aacaccgagg agaagggcga gatcaagaac tgctccttca acatcaccac cgagatccgc 480
gacaagaagc agaaggtgta cgccctgttc taccgcctgg acgtggtgcc catcgacgac 540
aacaacaaca actcctccaa ctaccgctg atcaactgca acacctccgc catcaccag 600
gcctgcccc aggtgtcctt cgagcccatc cccatccact actgcgcccc cgccggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcaccg gccctgcaa gaacgtgtcc 720
accgtgcagt gcaccacgg catcaagccc gtgggtgtcca ccagctgct gctgaacggc 780
tccctggccg aggaggagat catcatccgc tccgagaaca tcaccaacaa cgccaagacc 840
atcatcgtgc agctgaacga gtccgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagtcca tccgcatcgg ccccgccag gccttctacg ccaccggcga catcatcggc 960

gacatccgcc aggccactg caacatctcc ggcaccaagt ggaacaagac cctgcagcag 1020
 gtggccaaga agctgcgcga gcacttcaac aacaagacca tcatcttcaa gccctcctcc 1080
 ggcggcgacc tggagatcac caccactcc ttcaactgcc gcggcgagtt cttctactgc 1140
 aacacctccg gcctgttcaa ctccacctgg atcggaacg gcaccaagaa caacaacaac 1200
 accaacgaca ccatcacctt gccctgccg atcaagcaga tcatcaacat gtggcagggc 1260
 gtggggcagg ccatgtacgc ccccccatc gagggcaaga tcacctgcaa gtccaacatc 1320
 accggcctgc tgctgaccgc cgacggcggc aacaacaaca ccaacgagac cgagatcttc 1380
 cgccccggcg gcggcgacat gcgcgacaac tggcgctccg agctgtacaa gtacaagggtg 1440
 gtgaagatcg agccctggg cgtggcccc accaaggcca agctgaccgt gcaggcccg 1500
 cagctgctgt ccggcatcgt gcagcagcag tccaacctgc tgcgcgccat cgaggcccag 1560
 cagcacctgc tgcagctgac cgtgtggggc atcaagcagc tgcaggcccg cgtgctggcc 1620
 gtggagcgct acctgaagga ccagcagctg ctggagatct gggacaacat gacctggatg 1680
 gagtgggagc gcgagatcaa caactacacc gacatcatct actccctgat cgaggagtcc 1740
 cagaaccagc aggagaagaa cgagcaggag ctgtgtggcc tggacaagtg ggctccctg 1800
 tggaactggt tcgacatcac caactggctg tggtagatca agatcttcat catgatcgtg 1860
 ggcggcctga tcggcctgcg catcgtgttc gccgtgctgt ccatcgtgaa ccgcgtgcgc 1920
 cagggtact cccccctgtc cttccagacc ctgatcccca acccccgcg ccgaccgc 1980
 ccgaggggca tcgaggagga gggcgcgag caggaccgcg accgctccat ccgcctggtg 2040
 aacggcttcc tggccctggc ctgggacgac ctgcgctccc tgtgcctgtt ctctaccac 2100
 cgcctgcgcg acttcatcct gatcgccgc cgacccgtgg agctgctggg ccgcaagggc 2160
 ctgcgcgcg gctgggaggc cctgaagtac ctgtggaacc tgctgcagta ctggggccag 2220
 gagctgaaga actccgccat ctccctgctg gacaccaccg ccacgcctg ggccgagggc 2280
 accgaccgcg tgatcgaggt ggtgcagcgc gcctgccgcg ccacctgaa catccccgc 2340
 cgcatccgcc agggcctgga gcgcgcctg ctgtaa 2376

<210> 2
 <211> 50
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> IF-ApaI-SpPDI.c

<400> 2
 tgcccaaatt tgcggggccc atggcgaaaa acgttgcgat tttcggtta

50

<210> 3
 <211> 49

<212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> SpPDI-HIV gp145.r

<400> 3
 cacaggttct cggcggcgaa gatctgagaa ggaaccaaca caagaagag 49

<210> 4
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> IF-SpPDI-gp145.c

<400> 4
 tctcagatct tcgccgccga gaacctgtgg gtgaccgtgt actac 45

<210> 5
 <211> 38
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> WtTM-gp145.r

<400> 5
 cctttacagc agggcgcgct ccaggccctg gcggatgc 38

<210> 6
 <211> 4155
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Caset biểu hiện số 995

<400> 6
 ttaattaagt cgacaagctt gcatgccggt caacatgggt gagcacgaca cacttgtcta 60
 ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga cttttcaaca 120
 aagggttaata tccggaacc tcctcggatt ccattgccca gctatctgtc actttattgt 180
 gaagatagtg gaaaaggaag gtggctccta caaatgcat cattgcgata aaggaaaggc 240
 catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac ccacgaggag 300
 catcgtggaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt gatgtgataa 360
 catggtggag cagcacacac ttgtctactc caaaaatata aaagatacag tctcagaaga 420
 ccaaagggca attgagactt ttcaacaaag ggtaatatcc ggaaacctcc tcggattcca 480
 ttgccagct atctgtcact ttattgtgaa gatagtggaa aaggaagggt gctcctacaa 540
 atgccatcat tgcgataaag gaaaggccat cgttgaagat gcctctgccg acagtgggtc 600
 caaagatgga cccccacca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc 660

ttcaaagcaa gtggattgat gtgatatctc cactgacgta agggatgacg cacaatccca	720
ctatccttcg caagaccott cctctatata aggaagttca tttcatttgg agaggatta	780
aaatcttaat aggttttgat aaaagcgaac gtggggaaac ccgaaccaa ccttcttcta	840
aactctctct catctctctt aaagcaaact tctctcttgt ctttcttgcg tgagcgatct	900
tcaacgttgt cagatcgtgc ttcggcacca gtacaacgtt ttcttttact gaagcgaaat	960
caaagatctc tttgtggaca cgtagtgcg cgccattaaa taacgtgtac ttgtcctatt	1020
cttgctcggcg tggctctggg aaaagaaagc ttgctggagg ctgctgttca gcccataca	1080
ttacttgta cgattctgct gactttcggc gggtgcaata tctctacttc tgcttgacga	1140
ggtattgttg cctgtacttc tttcttcttc ttcttgctga ttggttctat aagaaatcta	1200
gtattttctt tgaaacagag ttttcccgtg gttttcgaac ttggagaaag attgttaagc	1260
ttctgtatat tctgcccaa tttgtcgggc ccatggcgaa aaacgttgcg attttcggct	1320
tattgttttc tcttcttggt ttggttcctt ctgagatctt cgccgccgag aacctgtggg	1380
tgaccgtgta ctacggcggtg cccgtgtgga aggaggccaa caccaccctg ttctgcgcct	1440
ccgacgccaa ggcctacgac accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc	1500
ccaccgacc caacccccag gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt	1560
ggaagaacaa catggtggag cagatgcacg aggacatcat ctccctgtgg gaccagtccc	1620
tgaagccctg cgtgaagctg acccccctgt gcgtgacct gaactgcacc aacgtgaacg	1680
tgaccaacac caccaacaac accgaggaga agggcgagat caagaactgc tccttcaaca	1740
tcaccaccga gatccgcgac aagaagcaga aggtgtacgc cctgttctac cgctggacg	1800
tggtgcccat cgacgacaac aacaacaact cctccaacta ccgcctgatc aactgcaaca	1860
cctccgccat caccaggcc tgccccaagg tgtccttcga gcccatcccc atccactact	1920
gcgccccgc cggttcgcc atcctgaagt gcaacgacaa gaagttcaac ggcaccggcc	1980
cctgcaagaa cgtgtccacc gtgcagtgc cccacggcat caagcccgtg gtgtccaccc	2040
agctgctgct gaacggctcc ctggccgagg aggagatcat catccgctcc gagaacatca	2100
ccaacaacgc caagaccatc atcgtgcagc tgaacgagtc cgtggagatc aactgcaccc	2160
gcccacaaca caacaccgc aagtccatcc gcacggccc cgccaggcc ttctacgcca	2220
ccggcgacat catcggcgac atccgccagg cccactgcaa catctccggc accaagtgga	2280
acaagacct gcagcaggtg gccagaagc tgcgcgagca cttcaacaac aagaccatca	2340
tcttcaagcc ctctccggc ggcgacctg agatcaccac ccactccttc aactgccgag	2400
gcgagttctt ctactgcaac acctccggc tgttcaactc cacctggatc ggcaacggca	2460
ccaagaacaa caacaacacc aacgacacca tcacctgcc ctgccgcatc aagcagatca	2520
tcaacatgtg gcaggcggtg ggccaggcca tgtacgcccc ccccatcgag ggcaagatca	2580

cctgcaagtc caacatcacc ggctgtgtgc tgacccgcga cggcggcaac aacaacacca 2640
 acgagaccga gatcttccgc cccggcggcg gcgacatgcg cgacaactgg cgctccgagc 2700
 tgtacaagta caaggtggtg aagatcgagc ccctgggctg ggcgccacc aaggccaagc 2760
 tgaccgtgca ggcccgccag ctgctgtccg gcatcggtga gcagcagtc aacctgtgtc 2820
 gcgccatcga ggcccagcag cacctgtgtc agctgaccgt gtggggcatc aagcagctgc 2880
 agggccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgtgt gagatctggg 2940
 acaacatgac ctggatggag tgggagcgcg agatcaacaa ctacaccgac atcatctact 3000
 ccctgatcga ggagtccag aaccagcagg agaagaacga gcaggagctg ctggccctgg 3060
 acaagtgggc ctccctgtgg aactggttcg acatcaccaa ctggctgtgg tacatcaaga 3120
 tcttcatcat gatcgtgggc ggctgatcg gcctgcgcat cgtgttcgcc gtgctgtcca 3180
 tcgtgaaccg cgtgcgccag ggctactccc ccctgtcctt ccagaccctg atccccaacc 3240
 cccgcggccc cgaccgcccc gagggcatcg aggaggaggg cggcgagcag gaccgcgacc 3300
 gctccatccg cctggtgaac ggcttcctgg ccctggcctg ggacgacctg cgctccctgt 3360
 gcctgttctc ctaccaccgc ctgcgcgact tcatcctgat cgccgcccgc accgtggagc 3420
 tgctgggccc caagggcctg cgccgcggct gggaggccct gaagtacctg tggaacctgc 3480
 tgcagtactg gggccaggag ctgaagaact ccgccatctc cctgctggac accaccgcca 3540
 tcgccgtggc cgagggcacc gaccgcgtga tcgaggtggt gcagcgcgcc tgccgcgcca 3600
 tcctgaacat ccccgccgc atccgccagg gcctggagcg cgccctgctg taaaggccta 3660
 ttttcttttag tttgaattta ctgttattcg gtgtgcattt ctatgtttgg tgagcggttt 3720
 tctgtgtcga gagtgtgttt attttatgta atttaatttc tttgtgagct cctgttttagc 3780
 aggtcgtccc ttacgaagg acacaaaaag attttaattt tattaataaaa aaaaaaaaaa 3840
 aagaccgga attcgatatc aagcttatcg acctgcagat cgttcaaaca tttggcaata 3900
 aagtttctta agattgaatc ctgttgccgg tcttgcgatg attatcatat aatttctgtt 3960
 gaattacgtt aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt 4020
 ttttatgatt agagtccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg 4080
 cgcaaactag gataaattat cgcgcgcggg gtcatttatg ttactagatc tctagagtct 4140
 caagcttggc gcgcc 4155

<210> 7

<211> 772

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của PDISP-HIV ConS (denta)CFI

<400> 7

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 Ala Glu Asn Leu Trp Val Thr Val
 20 25 30

Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Asn Thr Thr Leu Phe Cys
 35 40 45

Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala
 50 55 60

Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu
 65 70 75 80

Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met Val Glu
 85 90 95

Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro
 100 105 110

Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys Thr Asn Val
 115 120 125

Asn Val Thr Asn Thr Thr Asn Asn Thr Glu Glu Lys Gly Glu Ile Lys
 130 135 140

Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp Lys Lys Gln Lys
 145 150 155 160

Val Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val Pro Ile Asp Asp Asn
 165 170 175

Asn Asn Asn Ser Ser Asn Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala
 180 185 190

Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His
 195 200 205

Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys
 210 215 220

Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr
 225 230 235 240

His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser
 245 250 255

Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Ile Thr Asn Asn
 260 265 270

Ala Lys Thr Ile Ile Val Gln Leu Asn Glu Ser Val Glu Ile Asn Cys
 275 280 285

Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly
 290 295 300

Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala
 305 310 315 320

His Cys Asn Ile Ser Gly Thr Lys Trp Asn Lys Thr Leu Gln Gln Val
 325 330 335

Ala Lys Lys Leu Arg Glu His Phe Asn Asn Lys Thr Ile Ile Phe Lys
 340 345 350

Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys
 355 360 365

Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Ser Thr
 370 375 380

Trp Ile Gly Asn Gly Thr Lys Asn Asn Asn Asn Thr Asn Asp Thr Ile
 385 390 395 400

Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Gly Val
 405 410 415

Gly Gln Ala Met Tyr Ala Pro Pro Ile Glu Gly Lys Ile Thr Cys Lys
 420 425 430

Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn
 435 440 445

Thr Asn Glu Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp
 450 455 460

Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro
 465 470 475 480

Leu Gly Val Ala Pro Thr Lys Ala Lys Leu Thr Val Gln Ala Arg Gln
 485 490 495

Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile
 500 505 510

Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln
515 520 525

Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln
530 535 540

Leu Leu Glu Ile Trp Asp Asn Met Thr Trp Met Glu Trp Glu Arg Glu
545 550 555 560

Ile Asn Asn Tyr Thr Asp Ile Ile Tyr Ser Leu Ile Glu Glu Ser Gln
565 570 575

Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Ala Leu Asp Lys Trp
580 585 590

Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile
595 600 605

Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Val
610 615 620

Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro
625 630 635 640

Leu Ser Phe Gln Thr Leu Ile Pro Asn Pro Arg Gly Pro Asp Arg Pro
645 650 655

Glu Gly Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg Ser Ile
660 665 670

Arg Leu Val Asn Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu Arg Ser
675 680 685

Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile Ala
690 695 700

Ala Arg Thr Val Glu Leu Leu Gly Arg Lys Gly Leu Arg Arg Gly Trp
705 710 715 720

Glu Ala Leu Lys Tyr Leu Trp Asn Leu Leu Gln Tyr Trp Gly Gln Glu
725 730 735

Leu Lys Asn Ser Ala Ile Ser Leu Leu Asp Thr Thr Ala Ile Ala Val
740 745 750

Ala Glu Gly Thr Asp Arg Val Ile Glu Val Val Gln Arg Ala Cys Arg
755 760 765

Ala Ile Leu Asn
770

<210> 8
<211> 46
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF-H3dTm+gp145.r

<400> 8
ccatagtatc caatcgatgt accacagcca gttggtgatg tcgaac 46

<210> 9
<211> 49
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Gp145+H3TM.c

<400> 9
tggctgtggt acatcgattg gatactatgg atttcctttg ccatacat 49

<210> 10
<211> 35
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> H3dTm.r

<400> 10
ccttcaaag caaatgttgc acctaattgt gcctt 35

<210> 11
<211> 3735
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Caset biểu hiện số 997

<400> 11
ttaattaagt cgacaagctt gcatgccggt caacatggtg gagcacgaca cacttgtcta 60
ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga cttttcaaca 120
aagggttaata tccggaaacc tcctcggatt ccattgcccc gctatctgtc actttattgt 180
gaagatagtg gaaaaggaag gtggctccta caaatgccat cattgcgata aaggaaaggc 240
catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac ccacgaggag 300
catcgttgaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt gatgtgataa 360
catggtggag cacgacacac ttgtctactc caaaaatatc aaagatacag tctcagaaga 420

ccaaagggca attgagactt ttcaacaaag ggtaatatcc ggaaacctcc tcggattcca 480
 ttgcccagct atctgtcact ttattgtgaa gatagtggaa aaggaagggtg gtcctacaa 540
 atgccatcat tgcgataaag gaaaggccat cgttgaagat gcctctgccg acagtgggtcc 600
 caaagatgga cccccaccca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc 660
 ttcaaagcaa gtggattgat gtgatatctc cactgacgta agggatgacg cacaatccca 720
 ctatccttcg caagaccctt cctctatata aggaagttca tttcatttgg agaggtatta 780
 aaatcttaat aggttttggat aaaagcgaac gtggggaaac ccgaaccaa cttctttcta 840
 aactctctct catctctctt aaagcaaact tctctcttgt ctttcttgcg tgagcgatct 900
 tcaacgttgt cagatcgtgc ttcggcacca gtacaacgtt ttctttcact gaagcgaaat 960
 caaagatctc tttgtggaca cgtagtgcgg cgccattaaa taacgtgtac ttgtcctatt 1020
 cttgtcggtg tggctctggg aaaagaaagc ttgctggagg ctgctgttca gcccataca 1080
 ttacttggtta cgattctgct gactttcggc gggtgcaata tctctacttc tgcttgacga 1140
 ggtattgttg cctgtacttc tttcttcttc ttcttgctga ttggttctat aagaaatcta 1200
 gtattttctt tgaacagag ttttcccggt gttttcgaac ttggagaaaag attgttaagc 1260
 ttctgtatat tctgccaaa tttgtcgggc ccatggcgaa aaacgttgcg attttcggct 1320
 tattgttttc tcttcttggt ttggttcctt ctgagatctt cgccgccgag aacctgtggg 1380
 tgaccgtgta ctacggcgtg cccgtgtgga aggaggccaa caccaccctg ttctgcgctt 1440
 ccgacgcaa ggcctacgac accgaggtgc acaacgtgtg ggccaccac gcctgctgc 1500
 ccaccgaccc caacccccag gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt 1560
 ggaagaacaa catggtggag cagatgcacg aggacatcat ctccctgtgg gaccagtccc 1620
 tgaagccctg cgtgaagctg acccccctgt gcgtgacct gaactgcacc aacgtgaacg 1680
 tgaccaacac caccaacaac accgaggaga agggcgagat caagaactgc tccttcaaca 1740
 tcaccaccga gatccgcgac aagaagcaga aggtgtacgc cctgttctac cgcttgacg 1800
 tggtgcccat cgacgacaac aacaacaact cctccaacta ccgcctgac aactgcaaca 1860
 cctccgccat caccaggcc tgccccaagg tgtccttcga gcccatcccc atccactact 1920
 gcgccccgc cggttcgcc atcctgaagt gcaacgacaa gaagttcaac ggcaccggcc 1980
 cctgcaagaa cgtgtccacc gtgcagtga cccacggcat caagcccgtg gtgtccaccc 2040
 agctgctgct gaacggctcc ctggccgagg aggagatcat catccgctcc gagaacatca 2100
 ccaacaacgc caagaccatc atcgtgcagc tgaacgagtc cgtggagatc aactgcaccc 2160
 gcccacaaa caacaccgc aagtccatcc gcacggccc cgccaggcc ttctacgcca 2220
 ccggcgacat catcggcgac atccgccagg cccactgcaa catctccggc accaagtgga 2280
 acaagaccct gcagcaggtg gccagaagc tgcgcgagca cttaacaac aagaccatca 2340

```

tcttcaagcc ctctccggc ggcgacctg agatcaccac ccactccttc aactgccgcg      2400
gcgagttctt ctactgcaac acctccggcc tgttcaactc cacctggatc ggcaacggca      2460
ccaagaacaa caacaacacc aacgacacca tcacctgcc ctgccgcatc aagcagatca      2520
tcaacatgtg gcagggcggtg ggccaggcca tgtacgcccc ccccatcgag ggcaagatca      2580
cctgcaagtc caacatcacc ggctgtgtgc tgacctcgca cggcggcaac aacaacacca      2640
acgagaccga gatcttccgc cccggcggtg gcgacatgcg cgacaactgg cgctccgagc      2700
tgtacaagta caaggtggtg aagatcgagc ccctgggctg ggccccacc aaggccaagc      2760
tgacctgca ggcccgcag ctgctgtccg gcatcgatgca gcagcagtcc aacctgtgtc      2820
gcgccatcga ggcccagcag cacctgtgtc agctgacctg gtggggcatc aagcagctgc      2880
aggcccgctg gctggccgtg gagcgctacc tgaaggacca gcagctgtgt gagatctggg      2940
acaacatgac ctggatggag tgggagcgcg agatcaacaa ctacaccgac atcatctact      3000
ccctgatcga ggagtccag aaccagcagg agaagaacga gcaggagctg ctggccctgg      3060
acaagtgggc ctccctgtgg aactggttcg acatcaccaa ctggctgtgg tacatcgatt      3120
ggatactatg gatttccttt gccatatcat gttttttgct ttgtgttgct ttgttggggt      3180
tcatcatgtg ggctgcaa aaaggcaaca ttaggtgcaa catttgcat tgaaggccta      3240
ttttctttag tttgaattta ctgttattcg gtgtgcattt ctatgtttgg tgagcggttt      3300
tctgtgctca gagtgtgttt attttatgta atttaatttc tttgtgagct cctgttttagc      3360
aggtcgtccc ttcagcaagg acacaaaaag attttaattt tattaaaaaa aaaaaaaaaa      3420
aagaccggga attcgatatc aagcttatcg acctgcagat cgttcaaaca tttggcaata      3480
aagtttctta agattgaatc ctgttgccgg tcttgcatg attatcatat aatttctgtt      3540
gaattacgtt aagcatgtaa taattaacat gtaatgatg acgttattta tgagatgggt      3600
ttttatgatt agagtccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg      3660
cgcaaactag gataaattat cgcgcgcggt gtcatttatg ttactagatc tctagagtct      3720
caagcttggc gcgcc                                         3735

```

<210> 12

<211> 646

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của PDISP-HIV ConS
(denta)CFI-A/Brisbane/10/2007 H3 TM+CY

<400> 12

```

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 Ala Glu Asn Leu Trp Val Thr Val
 20 25 30

Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Asn Thr Thr Leu Phe Cys
 35 40 45

Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala
 50 55 60

Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu
 65 70 75 80

Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met Val Glu
 85 90 95

Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro
 100 105 110

Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys Thr Asn Val
 115 120 125

Asn Val Thr Asn Thr Thr Asn Asn Thr Glu Glu Lys Gly Glu Ile Lys
 130 135 140

Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp Lys Lys Gln Lys
 145 150 155 160

Val Tyr Ala Leu Phe Tyr Arg Leu Asp Val Val Pro Ile Asp Asp Asn
 165 170 175

Asn Asn Asn Ser Ser Asn Tyr Arg Leu Ile Asn Cys Asn Thr Ser Ala
 180 185 190

Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His
 195 200 205

Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys
 210 215 220

Phe Asn Gly Thr Gly Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr
 225 230 235 240

His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser
 245 250 255

Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Ile Thr Asn Asn
 260 265 270

Ala Lys Thr Ile Ile Val Gln Leu Asn Glu Ser Val Glu Ile Asn Cys
275 280 285

Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly
290 295 300

Gln Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala
305 310 315 320

His Cys Asn Ile Ser Gly Thr Lys Trp Asn Lys Thr Leu Gln Gln Val
325 330 335

Ala Lys Lys Leu Arg Glu His Phe Asn Asn Lys Thr Ile Ile Phe Lys
340 345 350

Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys
355 360 365

Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Ser Thr
370 375 380

Trp Ile Gly Asn Gly Thr Lys Asn Asn Asn Asn Thr Asn Asp Thr Ile
385 390 395 400

Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Gly Val
405 410 415

Gly Gln Ala Met Tyr Ala Pro Pro Ile Glu Gly Lys Ile Thr Cys Lys
420 425 430

Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn
435 440 445

Thr Asn Glu Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp
450 455 460

Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro
465 470 475 480

Leu Gly Val Ala Pro Thr Lys Ala Lys Leu Thr Val Gln Ala Arg Gln
485 490 495

Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile
500 505 510

Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln
515 520 525

Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln
 530 535 540

Leu Leu Glu Ile Trp Asp Asn Met Thr Trp Met Glu Trp Glu Arg Glu
 545 550 555 560

Ile Asn Asn Tyr Thr Asp Ile Ile Tyr Ser Leu Ile Glu Glu Ser Gln
 565 570 575

Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Ala Leu Asp Lys Trp
 580 585 590

Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile
 595 600 605

Asp Trp Ile Leu Trp Ile Ser Phe Ala Ile Ser Cys Phe Leu Leu Cys
 610 615 620

Val Ala Leu Leu Gly Phe Ile Met Trp Ala Cys Gln Lys Gly Asn Ile
 625 630 635 640

Arg Cys Asn Ile Cys Ile
 645

<210> 13
 <211> 46
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> IF-H5TM+gp145.r

<400> 13
 aattgacagt atttggatgt accacagcca gttggtgatg tcgaac

46

<210> 14
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Gp145+H5TM.c

<400> 14
 ctggctgtgg tacatccaaa tactgtcaat ttattcaaca gtggc

45

<210> 15
 <211> 35
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> IF-H5TM.r

<400> 15
cctttaaatg caaattctgc attgtaacga tccat

35

<210> 16
<211> 3735
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Caset biểu hiện số 999

<400> 16
ttaattaagt cgacaagctt gcatgccggt caacatgggt gagcacgaca cacttgtcta 60
ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga cttttcaaca 120
aagggttaata tccggaaacc tcctcggatt ccattgccca gctatctgtc actttattgt 180
gaagatagtg gaaaaggaag gtggctccta caaatgcat cattgcgata aaggaaaggc 240
catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac ccacgaggag 300
catcgtggaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt gatgtgataa 360
catggtggag cagcacacac ttgtctactc caaaaatatc aaagatacag tctcagaaga 420
ccaaagggca attgagactt ttcaacaaag ggtaatatcc ggaaacctcc tcggattcca 480
ttgccagct atctgtcact ttattgtgaa gatagtggaa aaggaagggt gtcctacaa 540
atgccatcat tgcgataaag gaaaggccat cgttgaagat gcctctgccg acagtgggcc 600
caaagatgga cccccacca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc 660
ttcaaagcaa gtggattgat gtgatatctc cactgacgta agggatgacg cacaatccca 720
ctatccttcg caagaccctt cctctatata aggaagtcca tttcatttgg agaggatta 780
aatctttaat aggttttcat aaaagcgaac gtggggaaac ccgaaccaa cttcttctc 840
aactctctct catctctctt aaagcaaact tctctcttgt ctttcttgcg tgagcgatct 900
tcaacgttgt cagatcgtgc ttcggcacca gtacaacgtt ttctttcact gaagcgaaat 960
caaagatctc tttgtggaca cgtagtgcgg cgccattaaa taacgtgtac ttgtcctatt 1020
cttgtcgggtg tggctctggg aaaagaaagc ttgctggagg ctgctgttca gcccataca 1080
ttacttggtta cgattctgct gactttcggc gggtgcaata tctctacttc tgcttgacga 1140
ggatttggtt cctgtacttc tttcttcttc ttcttgctga ttggttctat aagaaatcta 1200
gtattttctt tgaaacagag ttttcccggt gttttcgaac ttggagaaaag attgttaagc 1260
ttctgtatat tctgccccaa tttgtcgggc ccatggcgaa aaacgttgcg attttcggct 1320
tattgttttc tcttcttggt ttggttcctt ctcatatctt cgccgcccag aacctgtggg 1380
tgaccgtgta ctacggcggt cccgtgtgga aggaggccaa caccacctg ttctgcgcct 1440
ccgacgccaa ggcctacgac accgaggtgc acaacgtgtg ggccaccac gcctgcgtgc 1500

ccaccgaccc caacccccag gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt	1560
ggaagaacaa catggtggag cagatgcacg aggacatcat ctccctgtgg gaccagtccc	1620
tgaagccctg cgtgaagctg acccccctgt gcgtagaccct gaactgcacc aacgtgaacg	1680
tgaccaacac caccaacaac accgaggaga agggcgagat caagaactgc tccttcaaca	1740
tcaccaccga gatccgcgac aagaagcaga aggtgtacgc cctgttctac cgctggacg	1800
tggtgcccac cgacgacaac aacaacaact cctccaacta ccgcctgata aactgcaaca	1860
cctccgccat caccaggcc tgccccaagg tgtccttcga gcccatcccc atccactact	1920
gcgccccgc cggttcgac atcctgaagt gcaacgacaa gaagttcaac ggcaccggcc	1980
cctgcaagaa cgtgtccacc gtgcagtga cccacggcat caagcccgtg gtgtccaccc	2040
agctgctgct gaacggctcc ctggccgagg aggagatcat catccgctcc gagaacatca	2100
ccaacaacgc caagaccatc atcgtgcagc tgaacgagtc cgtggagatc aactgcaccc	2160
gccccaaaca caacaccgc aagtccatcc gcatcggccc cggccaggcc ttctacgcca	2220
ccggcgacat catcggcgac atccgccagg cccactgcaa catctccggc accaagtgga	2280
acaagaccct gcagcagggt gccaagaagc tgcgcgagca cttcaacaac aagaccatca	2340
tcttcaagcc ctctccggc ggcgacctgg agatcaccac ccactcctc aactgccgcg	2400
gcgagttctt ctactgcaac acctccggcc tgttcaactc cacctggatc ggcaacggca	2460
ccaagaacaa caacaacacc aacgacacca tcacctgcc ctgccgcatc aagcagatca	2520
tcaacatgtg gcagggcgtg ggccaggcca tgtacgcccc ccccatcgag ggcaagatca	2580
cctgcaagtc caacatcacc ggcctgctgc tgaccgcga cggcggcaac aacaacacca	2640
acgagaccga gatcttccgc cccggcggcg gcgacatgcg cgacaactgg cgctccgagc	2700
tgtacaagta caaggtggtg aagatcgagc ccctgggctg ggccccacc aaggccaagc	2760
tgaccgtgca ggcccgcag ctgctgtccg gcatcgtgca gcagcagtc aacctgctgc	2820
gcgccatcga ggcccagcag cacctgctgc agctgaccgt gtggggcatc aagcagctgc	2880
aggcccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg gagatctggg	2940
acaacatgac ctggatggag tgggagcgcg agatcaacaa ctacaccgac atcatctact	3000
ccctgatcga ggagtccag aaccagcagg agaagaacga gcaggagctg ctggccctgg	3060
acaagtgggc ctccctgtgg aactggttcg acatcaccaa ctggctgtgg tacatccaaa	3120
tactgtcaat ttattcaaca gtggcgagtt ccctagcact ggcaatcatg atggctggtc	3180
tatctttatg gatgtgctcc aatggatcgt tacaatgcag aatttgcat taaaggccta	3240
ttttcttttag tttgaattta ctgttattcg gtgtgcattt ctatgtttgg tgagcggttt	3300
tctgtgotca gagtgtgttt attttatgta atttaatttc tttgtgagct cctgttttagc	3360
aggctgtccc ttcagcaagg acacaaaaag attttaattt tattaaaaaa aaaaaaaaaa	3420

```

aagaccggga attcgatatc aagcttatcg acctgcagat cgttcaaaca tttggcaata      3480
aagttttotta agattgaatc ctgttgccgg tcttgcgatg attatcatat aattttctggt      3540
gaattacggtt aagcatgtaa taattaacat gtaatgcatg acgttattta tgagatgggt      3600
ttttatgatt agagtcccg c aattatacat ttaatacgcg atagaaaaca aaatatagcg      3660
cgcaaactag gataaattat cgcgcgcggg gtcactctatg ttactagatc tctagagtct      3720
caagcttggc gcgcc                                                    3735

```

<210> 17
 <211> 646
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Trình tự axit amino của PDISP-HIV ConS
 (denta)CFI-A/Indonesia/5/2005 H5 TM+CY

<400> 17

```

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 Ala Glu Asn Leu Trp Val Thr Val
              20              25              30

```

```

Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Asn Thr Thr Leu Phe Cys
              35              40              45

```

```

Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala
              50              55              60

```

```

Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu
65              70              75              80

```

```

Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asn Met Val Glu
              85              90              95

```

```

Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro
              100              105              110

```

```

Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu Asn Cys Thr Asn Val
              115              120              125

```

```

Asn Val Thr Asn Thr Thr Asn Asn Thr Glu Glu Lys Gly Glu Ile Lys
              130              135              140

```

```

Asn Cys Ser Phe Asn Ile Thr Thr Glu Ile Arg Asp Lys Lys Gln Lys
145              150              155              160

```

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

Gly Gln Ala Met Tyr Ala Pro Pro Ile Glu Gly Lys Ile Thr Cys Lys
 420 425 430

Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn
 435 440 445

Thr Asn Glu Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp
 450 455 460

Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro
 465 470 475 480

Leu Gly Val Ala Pro Thr Lys Ala Lys Leu Thr Val Gln Ala Arg Gln
 485 490 495

Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile
 500 505 510

Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln
 515 520 525

Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln
 530 535 540

Leu Leu Glu Ile Trp Asp Asn Met Thr Trp Met Glu Trp Glu Arg Glu
 545 550 555 560

Ile Asn Asn Tyr Thr Asp Ile Ile Tyr Ser Leu Ile Glu Glu Ser Gln
 565 570 575

Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Ala Leu Asp Lys Trp
 580 585 590

Ala Ser Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile
 595 600 605

Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala
 610 615 620

Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu
 625 630 635 640

Gln Cys Arg Ile Cys Ile
 645

<210> 18
 <211> 1967
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Caset biểu hiện số 172

<400> 18

```

cccgggctgg tatatttata tgttgtcaaa taactcaaaa accataaaaag ttttaagttag      60
caagtgtgta catttttact tgaacaaaaa tattcaccta ctactgttat aaatcattat      120
taaacattag agtaaagaaa tatggatgat aagaacaaga gtagtgatat tttgacaaca      180
atthttgttg aacatttgag aaaattttgt tgttctctct tttcattggg caaaaacaat      240
agagagagaa aaaggaagag ggagaataaa aacataatgt gagtatgaga gagaaagttg      300
tacaaaagtt gtacccaaat agttgtacaa atatcattga ggaatttgac aaaagctaca      360
caaataaggg ttaattgctg taaataaata aggatgacgc attagagaga tgtaccatta      420
gagaattttt ggcaagtcac taaaaagaaa gaataaatta tttttaaaat taaaagttga      480
gtcatttgat taaacatgtg attattttaat gaattgatga aagagttgga ttaaagttgt      540
attagtaatt agaatttggt gtcaaattta atttgacatt tgatcttttc ctatatattg      600
ccccatagag tcagttaact catttttata tttcatagat caaataagag aaataacggt      660
atattaatcc ctccaaaaaa aaaaaacggt atatttacta aaaaatctaa gccacgtagg      720
aggataacag gatccccgta ggaggataac atccaatcca accaatcaca acaatcctga      780
tgagataacc cactttaagc ccacgcatct gtggcacatc tacattatct aaatcacaca      840
ttcttccaca catctgagcc acacaaaaac caatccacat ctttatcacc cattctataa      900
aaaatcacac tttgtgagtc tacactttga ttcccttcaa acacatacaa agagaagaga      960
ctaattaatt aattaatcat cttgagagaa aatggaacga gctatacaag gaaacgacgc     1020
tagggaacaa gctaacagtg aacgttgagg tggaggatca ggaggtagca cttctccctt     1080
caaacttcct gacgaaagtc cgagttggac tgagtggcgg ctacataacg atgagacgaa     1140
ttcgaatcaa gataatcccc ttggtttcaa ggaaagctgg ggtttcggga aagttgtatt     1200
taagagatat ctgagatacg acaggacgga agcttactg cacagagtcc ttggatcttg     1260
gacgggagat tcggttaact atgcagcatc tcgatttttc ggtttcgacc agatcggatg     1320
tacctatagt attcggtttc gaggagttag tatcaccggt tctggagggt cgcgaactct     1380
tcagcatctc tgtgagatgg caattcggtc taagcaagaa ctgctacagc ttgccccaat     1440
cgaagtggaa agtaatgtat caagaggatg ccctgaaggt actcaaacct tcgaaaaaga     1500
aagcgagtaa gtcgagggcg agctctaagt taaaatgctt cttcgtctcc tatttataat     1560
atggtttggt attgttaatt ttgttcttgt agaagagctt aattaatcgt tgttgttatg     1620
aaatactatt tgtatgagat gaactggtgt aatgtaattc atttacataa gtggagtcag     1680
aatcagaatg tttcctccat aactaactag acatgaagac ctgccgcgta caattgtctt     1740
atatttgaac aactaaaatt gaacatcttt tgccacaact ttataagtgg ttaatatagc     1800

```

tcaaatatat ggtcaagttc aatagattaa taatggaaat atcagttatc gaaattcatt 1860
 aacaatcaac ttaacgttat taactactaa ttttatatca tcccctttga taaatgatag 1920
 tacaccaatt aggaaggagc atgctcgagg cctggctggc cgaattc 1967

<210> 19
 <211> 172
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Trình tự axit amin của gen ức chế TBSV P19 bị bất hoạt

<400> 19

Met Glu Arg Ala Ile Gln Gly Asn Asp Ala Arg Glu Gln Ala Asn Ser
 1 5 10 15

Glu Arg Trp Asp Gly Gly Ser Gly Gly Thr Thr Ser Pro Phe Lys Leu
 20 25 30

Pro Asp Glu Ser Pro Ser Trp Thr Glu Trp Arg Leu His Asn Asp Glu
 35 40 45

Thr Asn Ser Asn Gln Asp Asn Pro Leu Gly Phe Lys Glu Ser Trp Gly
 50 55 60

Phe Gly Lys Val Val Phe Lys Arg Tyr Leu Arg Tyr Asp Arg Thr Glu
 65 70 75 80

Ala Ser Leu His Arg Val Leu Gly Ser Trp Thr Gly Asp Ser Val Asn
 85 90 95

Tyr Ala Ala Ser Arg Phe Phe Gly Phe Asp Gln Ile Gly Cys Thr Tyr
 100 105 110

Ser Ile Arg Phe Arg Gly Val Ser Ile Thr Val Ser Gly Gly Ser Arg
 115 120 125

Thr Leu Gln His Leu Cys Glu Met Ala Ile Arg Ser Lys Gln Glu Leu
 130 135 140

Leu Gln Leu Ala Pro Ile Glu Val Glu Ser Asn Val Ser Arg Gly Cys
 145 150 155 160

Pro Glu Gly Thr Gln Thr Phe Glu Lys Glu Ser Glu
 165 170

<210> 20
 <211> 47

<212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> IF-wtSp-VZVgE.c

<400> 20
 tgcccaaatt tgtcgggcc atggggacag ttaataaacc tgtggtg 47

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

<220>
 <223> IF-H5TM+VZVgE.r

<400> 21
 aattgacagt atttgtcgta gaagtgggtga cgttccgggg tttacg 46

<210> 22
 <211> 1872
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> synthesized VZV gE gene (corresponding to nt 3477-5348 from
 Genbank accession

number AY013752.1)

<400> 22
 atggggacag ttaataaacc tgtggtgggg gtattgatgg ggttcggaat tatcacggga 60
 acgttgcgta taacgaatcc ggtcagagca tccgtcttgc gatacgatga ttttcacacc 120
 gatgaagaca aactggatac aaactccgta tatgagcctt actaccattc agatcatgcg 180
 gagtcttcat gggtaaactcg gggagagtct tcgcgaaaag cgtacgatca taactcacct 240
 tatatatggc cacgtaatga ttatgatgga tttttagaga acgcacacga acaccatggg 300
 gtgtataatc agggccgtgg tatcgatagc ggggaacggt taatgcaacc cacacaaatg 360
 tctgcacagg aggatcttgg ggacgatacg ggcacccacg ttatccctac gttaaacggc 420
 gatgacagac ataaaattgt aaatgtggac caacgtcaat acggtgacgt gtttaaagga 480
 gatcttaatc caaaacccca aggccaaaga ctcatgagg tgtcagtgga agaaaatcac 540
 ccgtttactt tacgcgcacc gattcagcgg atttatggag tccggtacac cgagacttgg 600
 agctttttgc cgtcattaac ctgtacggga gacgcagcgc ccgccatcca gcatatatgt 660
 ttaaaacata caacatgctt tcaagacgtg gtggtggatg tggattgcgc ggaaaatact 720
 aaagaggatc agttggccga aatcagttac cgttttcaag gtaagaagga agcggaccaa 780
 ccgtggattg ttgtaaacac gagcacttg tttgatgaac tcgaattaga ccccccgag 840
 attgaaccgg gtgtcttgaa agtacttcgg acagaaaaac aatacttggg tgtgtacatt 900

tggaacatgc gcggctccga tggtagctct acctacgccg cgtttttggt cacctggaaa 960
 ggggatgaaa aaacaagaaa ccctacgccc gcagtaactc ctcaaccaag aggggctgag 1020
 ttccatattgt ggaattacca ctgcgatgta ttttcagttg gtgatacgtt tagcttggca 1080
 atgcattcttc agtataagat acatgaagcg ccatttgatt tgctgttaga gtgggtgtat 1140
 gtccccatcg atcctacatg tcaaccaatg cggttatatt ctacgtgttt gtatcatccc 1200
 aacgcacccc aatgcctctc tcatatgaat tccggttgta catttacctc gccacattta 1260
 gccagcgtg ttgcaagcac agtgtatcaa aattgtgaac atgcagataa ctacaccgca 1320
 tattgtctgg gaatatctca tatggagcct agctttggtc taatcttaca cgacgggggc 1380
 accacgttaa agttttaga tacacccgag agtttgcgg gattatacgt ttttgtggtg 1440
 tatttttaacg ggcatgttg agccgtagca tacactgttg tatccacagt agatcatttt 1500
 gtaaacgcaa ttgaagagcg tggatttccg ccaacggccg gtcagccacc ggcgactact 1560
 aaaccaagg aaattacccc cgtaaaccgc ggaacgtcac cacttctacg atatgccgca 1620
 tggaccggag ggcttgacg agtagtactt ttatgtctcg taatattttt aatctgtacg 1680
 gctaaacgaa tgaggggttaa agcctatagg gtagacaagt ccccgataa ccaaagcatg 1740
 tattacgctg gccttccagt ggacgatttc gaggactcgg aatctacgga tacggaagaa 1800
 gagtttgata acgcgattgg agggagtcac gggggttcga gttacacggt gtatatagat 1860
 aagaccgggt ga 1872

<210> 23
 <211> 46
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> VZVgE+H5TM.c

<400> 23
 gccactgttg aataaattga cagtatttgc cgtagaagtg gtgacg

46

<210> 24
 <211> 3522
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Caset biểu hiện số 946

<400> 24
 ttaattaagt cgacaagctt gcatgccggt caacatggtg gagcacgaca cacttgtcta 60
 ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga cttttcaaca 120
 aagggttaata tccggaaacc tcctcggatt ccattgccca gctatctgtc actttattgt 180
 gaagatagtg gaaaaggaag gtggctccta caaatgccat cattgcgata aaggaaaggc 240

catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac ccacgaggag	300
catcgtggaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt gatgtgataa	360
catggtggag cacgacacac ttgtctactc caaaaatatc aaagatacag tctcagaaga	420
ccaaagggca attgagactt ttcaacaaaag ggtaatatcc ggaaacctcc tcggattcca	480
ttgccagct atctgtcact ttattgtgaa gatagtggaa aaggaagggtg gctcctacaa	540
atgccatcat tgcgataaag gaaaggccat cgttgaagat gcctctgccg acagtgggtcc	600
caaagatgga cccccacca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc	660
ttcaaagcaa gtggattgat gtgatatctc cactgacgta agggatgacg cacaatccca	720
ctatccttcg caagaccctt cctctatata aggaagttca tttcatttgg agaggтата	780
aaatcttaat aggttttgat aaaagcgaac gtggggaaaac ccgaaccaa ccttcttcta	840
aactctctct catctctctt aaagcaaact tctctcttgt ctttcttgcg tgagcgatct	900
tcaacgttgt cagatcgtgc ttcggcacca gtacaacgtt ttcttttcaact gaagcgaaat	960
caaagatctc tttgtggaca cgtagtgcgg cgccattaaa taacgtgtac ttgtcctatt	1020
cttgtcggtg tggctcttggg aaaagaaagc ttgctggagg ctgctgttca gcccataca	1080
ttacttgta cgattctgct gactttcggc gggtgcaata tctctacttc tgcttgacga	1140
ggtattgttg cctgtacttc tttcttcttc ttcttgctga ttggttctat aagaaatcta	1200
gtattttctt tgaaacagag ttttcccggt gttttcgaac ttggagaaaag attgttaagc	1260
ttctgtatat tctgccccaa tttgtcgggc ccatggggac agttaataaa cctgtggtgg	1320
gggtattgat ggggttcgga attatcacgg gaacgttgcg tataacgaat ccggtcagag	1380
catccgtctt gcgatacgat gattttcaca ccgatgaaga caaactggat acaaactccg	1440
tatatgagcc ttactacat tcagatcatg cggagtcttc atgggtaa at cggggagagt	1500
cttcgcgaaa agcgtacgat cataactcac cttatatatg gccacgtaat gattatgatg	1560
gattttttaga gaacgcacac gaacaccatg ggggtgataa tcagggccgt ggtatcgata	1620
gcggggaacg gttaatgcaa cccacacaaa tgtctgcaca ggaggatctt ggggacgata	1680
cgggcatcca cgttatccct acgttaaacg gcgatgacag acataaaatt gtaaatgtgg	1740
accaacgtca atacggtgac gtgttttaaag gagatcttaa tccaaaacc caaggccaaa	1800
gactcattga ggtgtcagtg gaagaaaatc acccgtttac tttacgcgca ccgattcagc	1860
ggatttatgg agtccggtac accgagactt ggagcttttt gccgtcatta acctgtacgg	1920
gagacgcagc gcccgccatc cagcatatat gtttaaaaca tacaacatgc tttcaagacg	1980
tgggtgggga tgtggattgc gcggaaaata ctaaagagga tcagttggcc gaaatcagtt	2040
accgttttca aggtagaag gaagcggacc aaccgtggat tgttgtaa ac acgagcacac	2100
tgtttgatga actcgaatta gaccccccg agattgaacc ggggtgtcttg aaagtacttc	2160

```

ggacagaaaa acaataacttg ggtgtgtaca tttggaacat gcgcggctcc gatggtacgt      2220
ctacctacgc cacgttttttg gtcacctgga aaggggatga aaaaacaaga aaccttacgc      2280
ccgcagtaac tcotcaacca agaggggctg agtttcatat gtggaattac cactcgcatg      2340
tattttcagt tggtgatagc tttagcttgg caatgcatct tcagtataag atacatgaag      2400
cgccatttga tttgctgtta gagtggttgt atgtcccat cgatcctaca tgtcaaccaa      2460
tgcggttata ttctacgtgt ttgtatcatc ccaacgcacc ccaatgcctc totcatatga      2520
attccggttg tacatttacc tcgccacatt tagcccagcg tgttgcaagc acagtgtatc      2580
aaaattgtga acatgcagat aactacaccg catattgtct gggaatatct catatggagc      2640
ctagcttttg tctaattcta cacgacgggg gcaccacgtt aaagtttgta gataacccg      2700
agagtttgtc gggattatac gtttttgtgg tgtattttta cgggcatgtt gaagccgtag      2760
catacactgt tgtatccaca gtagatcatt ttgtaaacgc aattgaagag cgtggatttc      2820
cgccaacggc cggtcagcca ccggcgacta ctaaacccaa ggaaattacc cccgtaaacc      2880
ccggaacgtc accacttcta cgacaaatac tgtcaattta ttcaacagtg gcgagttccc      2940
tagcactggc aatcatgatg gctggtctat ctttatggat gtgctccaat ggatcgttac      3000
aatgcagaat ttgcatttaa aggcctatct tctttagttt gaatttactg ttattcggtg      3060
tgcatttcta tgtttggtga gcggttttct gtgctcagag tgtgtttatt ttatgtaatt      3120
taatttcttt gtgagctcct gtttagcagg tcgtcccttc agcaaggaca caaaaagatt      3180
ttaattttat taaaaaaaaa aaaaaaaaaa accgggaatt cgatatcaag cttatcgacc      3240
tgcagatcgt tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct      3300
tgcgatgatt atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta      3360
atgcatgacg ttattttatga gatgggtttt tatgattaga gtcccgcaat tatacattta      3420
atacgcgata gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc      3480
atctatgtta ctagatctct agagtctcaa gcttggcgcg cc                          3522

```

<210> 25

<211> 575

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Amino acid sequence of VZV gE- A/Indonesia/5/2005 H5 TM+CY

<400> 25

```

Met Gly Thr Val Asn Lys Pro Val Val Gly Val Leu Met Gly Phe Gly
1           5           10          15

```

```

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

```

Leu Arg Tyr Asp Asp Phe His Thr Asp Glu Asp Lys Leu Asp Thr Asn
 35 40 45

Ser Val Tyr Glu Pro Tyr Tyr His Ser Asp His Ala Glu Ser Ser Trp
 50 55 60

Val Asn Arg Gly Glu Ser Ser Arg Lys Ala Tyr Asp His Asn Ser Pro
 65 70 75 80

Tyr Ile Trp Pro Arg Asn Asp Tyr Asp Gly Phe Leu Glu Asn Ala His
 85 90 95

Glu His His Gly Val Tyr Asn Gln Gly Arg Gly Ile Asp Ser Gly Glu
 100 105 110

Arg Leu Met Gln Pro Thr Gln Met Ser Ala Gln Glu Asp Leu Gly Asp
 115 120 125

Asp Thr Gly Ile His Val Ile Pro Thr Leu Asn Gly Asp Asp Arg His
 130 135 140

Lys Ile Val Asn Val Asp Gln Arg Gln Tyr Gly Asp Val Phe Lys Gly
 145 150 155 160

Asp Leu Asn Pro Lys Pro Gln Gly Gln Arg Leu Ile Glu Val Ser Val
 165 170 175

Glu Glu Asn His Pro Phe Thr Leu Arg Ala Pro Ile Gln Arg Ile Tyr
 180 185 190

Gly Val Arg Tyr Thr Glu Thr Trp Ser Phe Leu Pro Ser Leu Thr Cys
 195 200 205

Thr Gly Asp Ala Ala Pro Ala Ile Gln His Ile Cys Leu Lys His Thr
 210 215 220

Thr Cys Phe Gln Asp Val Val Val Asp Val Asp Cys Ala Glu Asn Thr
 225 230 235 240

Lys Glu Asp Gln Leu Ala Glu Ile Ser Tyr Arg Phe Gln Gly Lys Lys
 245 250 255

Glu Ala Asp Gln Pro Trp Ile Val Val Asn Thr Ser Thr Leu Phe Asp
 260 265 270

Glu Leu Glu Leu Asp Pro Pro Glu Ile Glu Pro Gly Val Leu Lys Val
 275 280 285

Leu Arg Thr Glu Lys Gln Tyr Leu Gly Val Tyr Ile Trp Asn Met Arg
 290 295 300

Gly Ser Asp Gly Thr Ser Thr Tyr Ala Thr Phe Leu Val Thr Trp Lys
 305 310 315 320

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

Arg Gly Ala Glu Phe His Met Trp Asn Tyr His Ser His Val Phe Ser
 340 345 350

Val Gly Asp Thr Phe Ser Leu Ala Met His Leu Gln Tyr Lys Ile His
 355 360 365

Glu Ala Pro Phe Asp Leu Leu Leu Glu Trp Leu Tyr Val Pro Ile Asp
 370 375 380

Pro Thr Cys Gln Pro Met Arg Leu Tyr Ser Thr Cys Leu Tyr His Pro
 385 390 395 400

Asn Ala Pro Gln Cys Leu Ser His Met Asn Ser Gly Cys Thr Phe Thr
 405 410 415

Ser Pro His Leu Ala Gln Arg Val Ala Ser Thr Val Tyr Gln Asn Cys
 420 425 430

Glu His Ala Asp Asn Tyr Thr Ala Tyr Cys Leu Gly Ile Ser His Met
 435 440 445

Glu Pro Ser Phe Gly Leu Ile Leu His Asp Gly Gly Thr Thr Leu Lys
 450 455 460

Phe Val Asp Thr Pro Glu Ser Leu Ser Gly Leu Tyr Val Phe Val Val
 465 470 475 480

Tyr Phe Asn Gly His Val Glu Ala Val Ala Tyr Thr Val Val Ser Thr
 485 490 495

Val Asp His Phe Val Asn Ala Ile Glu Glu Arg Gly Phe Pro Pro Thr
 500 505 510

Ala Gly Gln Pro Pro Ala Thr Thr Lys Pro Lys Glu Ile Thr Pro Val
 515 520 525

Asn Pro Gly Thr Ser Pro Leu Leu Arg 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> 26
<211> 50
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF-wtSp-SARSgS.c

<400> 26
tgcccaaatt tgcgggccc atgtttatit tcttattatt tcttactctc 50

<210> 27
<211> 45
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> IF-H5TM+SARSgS.r

<400> 27
aattgacagt atttgaggcc atttaatat ttgtcatat tttcc 45

<210> 28
<211> 3768
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> synthesized SARS gS gene (corresponding to nt 21492-25259 from
Genebank

accession number AY278741.1)

<400> 28
atgtttatit tcttattatt tcttactctc actagtggta gtgaccttga ccggtgcacc 60
acttttgatg atgttcaagc tcctaattac actcaacata cttcatctat gagggggggtt 120
tactatcctg atgaaatit tagatcagac actctttatt taactcagga tttatttctt 180
ccattttatt ctaatgttac agggtttcat actattaatc atacgttttg caaccctgtc 240
atacctttta aggatgggtat ttattttgct gccacagaga aatcaaagt tgtccgtgggt 300
tgggtttttg gttctacat gaacaacaag tcacagtcgg tgattattat taacaattct 360
actaatgttg ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctgtt 420
tctaaacca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
ttcgagtaca tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaaa 540
cacttacgag agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggtat 600

caacctatag atgtagttcg tgatctacct totggtttta acactttgaa acctatTTTT	660
aagttgcctc ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct	720
gctcaagaca tttggggcac gtcagctgca gcctatTTTg ttggctatTT aaagccaaact	780
acatttatgc tcaagtatga tgaaaatggg acaatcacag atgctgttga ttgtttctcaa	840
aatccacttg ctgaactcaa atgctctgtt aagagctttg agattgacaa aggaattttac	900
cagacctcta atttcagggt tgttccctca ggagatgttg tgagattccc taatattaca	960
aacttgtgtc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg	1020
gagagaaaaa aaattttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt	1080
ttttcaacct ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc	1140
aatgtctatg cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgccagga	1200
caaactgggtg ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc	1260
cttgcttgga atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat	1320
aggtatctta gacatggcaa gcttaggccc tttgagagag acatatctaa tgtgcctttc	1380
tcccctgatg gcaaaccttg caccacacct gctcttaatt gttattggcc attaaatgat	1440
tatggttttt acaccactac tggcattggc taccaacctt acagagttgt agtactttct	1500
tttgaacttt taaatgcacc ggccacgggt tgtggaccaa aattatccac tgaccttatt	1560
aagaaccagt gtgtcaattt taattttaat ggactcactg gtactgggtg gttaactcct	1620
tcttcaaaga gatttcaacc atttcaacaa tttggccgtg atgtttctga tttcactgat	1680
tccgttcgag atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt	1740
gtaagtgtaa ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat	1800
gttaactgca ctgatgtttc tacagcaatt catgcagatc aactcacacc agcttggcgc	1860
atatattcta ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag	1920
catgtcgaca cttcttatga gtgcgacatt cctattggag ctggcatttg tgctagttac	1980
catacagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct	2040
ttaggtgctg atagttcaat tgcttactct aataacacca ttgctatacc tactaacttt	2100
tcaattagca ttactacaga agtaatgcct gtttctatgg ctaaaacctc cgtagattgt	2160
aatatgtaca tctgcggaga ttctactgaa tgtgctaatt tgcttctcca atatggtagc	2220
ttttgcacac aactaaatcg tgcaactctc ggtattgctg ctgaacagga tcgcaacaca	2280
cgtgaagtgt tcgctcaagt caaacaaatg tacaaaaccc caactttgaa atatTTTtgg	2340
ggttttaatt tttcacaaat attacctgac cctctaaagc caactaagag gtcttttatt	2400
gaggacttgc tctttaataa ggtgacactc gctgatgctg gcttcatgaa gcaatatggc	2460
gaatgcctag gtgatattaa tgctagagat ctcatTTgtg cgcagaagtt caatggactt	2520

acagtgttgc cacctctgct cactgatgat atgattgctg cctacactgc tgctctagtt 2580
 agtgggtactg ccactgctgg atggacattt ggtgctggcg ctgctcttca aatacctttt 2640
 gctatgcaaa tggcatatag gttcaatggc attggagtta cccaaaatgt tctctatgag 2700
 aaccaaaaaac aaatcgccaa ccaatttaac aaggcgatta gtcaaattca agaatacatt 2760
 acaacaacat caactgcatt gggcaagctg caagacgttg ttaaccagaa tgctcaagca 2820
 ttaaacacac ttgttaaaca acttagctct aatttttggtg caatttcaag tgtgctaaat 2880
 gatatccttt cgcgacttga taaagtcgag gcggaggtac aaattgacag gttaattaca 2940
 ggcagacttc aaagccttca aacctatgta acacaacaac taatcagggc tgctgaaatc 3000
 agggcttctg ctaatcttgc tgctactaaa atgtctgagt gtgttcttgg acaatcaaaa 3060
 agagttgact tttgtggaaa gggctaccac cttatgtcct tcccacaagc agccccgcat 3120
 ggtgttgtct tcctacatgt cacgtatgtg ccatcccagg agaggaactt caccacagcg 3180
 ccagcaattt gtcataagg caaagcatac ttccctcgtg aagggtgtttt tgtgtttaat 3240
 ggcacttctt ggtttattac acagaggaac ttcttttctc cacaataat tactacagac 3300
 aatacatttg tctcaggaaa ttgtgatgtc gttattggca tcattaacaa cacagtttat 3360
 gatcctctgc aacctgagct cgactcattc aaagaagagc tggacaagta cttcaaaaat 3420
 catacatcac cagatgttga tcttggcgac atttcaggca ttaacgcttc tgctgtcaac 3480
 attcaaaaag aaattgaccg cctcaatgag gtcgctaaaa atttaaata atcactcatt 3540
 gaccttcaag aattgggaaa atatgagcaa tatattaaat ggccttggtg tgtttggtc 3600
 ggcttcattg ctggactaat tgccatcgtc atgggtacaa tcttgctttg ttgcatgact 3660
 agttgttgca gttgcctcaa ggggtgatgc tcttggtggt cttgctgcaa gtttgatgag 3720
 gatgactctg agccagttct caagggtgtc aaattacatt acacataa 3768

<210> 29
 <211> 43
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> SARSgS+H5TM.c

<400> 29
 atatattaaa tggcctcaaa tactgtcaat ttattcaaca gtg 43

<210> 30
 <211> 5496
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Caset biểu hiện số 916

```

<400> 30
ttaattaagt cgacaagctt gcatgccggt caacatggtg gagcacgaca cacttgtcta      60
ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga cttttcaaca      120
aagggttaata tccggaaaacc tcctcggatt ccattgcccga gctatctgtc actttattgt      180
gaagatagtg gaaaaggaag gtggctccta caaatgccat cattgcgata aaggaaaggc      240
catcgttgaa gatgcctctg ccgacagtgg toccaaagat ggacccccac ccacgaggag      300
catcgtggaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt gatgtgataa      360
catggtggag cacgacacac ttgtctactc caaaaatatc aaagatacag tctcagaaga      420
ccaaagggca attgagactt ttcaacaaaag ggtaatatcc ggaaacctcc tcggattcca      480
ttgccagct atctgtcact ttattgtgaa gatagtggaa aagggaagggtg gctcctacaa      540
atgccatcat tgcgataaag gaaaggccat cgttgaagat gcctctgccg acagtgggtcc      600
caaagatgga cccccacca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc      660
ttcaaagcaa gtggattgat gtgatctctc cactgacgta agggatgacg cacaatccca      720
ctatccttcg caagaccctt cctctatata aggaagttca tttcatttgg agaggtatta      780
aaatcttaat aggttttgat aaaagcgaac gtggggaaac ccgaaccaa ccttcttcta      840
aactctctct catctctctt aaagcaaact tctctcttgt ctttcttgcg tgagcgatct      900
tcaacgttgt cagatcgtgc ttcggcacca gtacaacgtt ttctttcact gaagcgaaat      960
caaagatctc tttgtggaca cgtagtgcgg cgccattaaa taacgtgtac ttgtcctatt      1020
cttgtcggtg tggctttggg aaaagaaagc ttgctggagg ctgctgttca gcccataca      1080
ttacttgta cgattctgct gactttcggc gggtgcaata tctctacttc tgcttgacga      1140
ggatttggtt cctgtacttc tttcttcttc ttcttgctga ttggttctat aagaaatcta      1200
gtattttctt tgaaacagag ttttcccggt gttttcgaac ttggagaaaag attgttaagc      1260
ttctgtatat tctgcccata tttgtcgggc ccatgtttat tttcttatta tttcttactc      1320
tcactagtgg tagtgacctt gaccggtgca ccacttttga tgatgttcaa gctcctaatt      1380
acactcaaca tacttcatct atgagggggg tttactatcc tgatgaaatt tttagatcag      1440
acactcttta tttaactcag gatttatctt ttccatttta ttctaattgt acagggtttc      1500
atactattaa tcatacgttt ggcaaccctg tcataccttt taaggatggg atttattttg      1560
ctgccacaga gaaatcaaat gttgtccgtg gttgggtttt tggttctacc atgaacaaca      1620
agtcacagtc ggtgattatt attaacaatt ctactaatgt tgttatacga gcatgtaact      1680
ttgaattgtg tgacaaccct ttctttgctg tttctaaacc catgggtaca cagacacata      1740
ctatgatatt cgataatgca tttaattgca ctttcgagta catatctgat gccttttcgc      1800
ttgatgtttc agaaaagtca ggtaatttta aacacttacg agagtttgtg tttaaaaata      1860
aagatggggt tctctatggt tataagggct atcaacctat agatgtagtt cgtgatctac      1920

```

cttctggttt taacactttg aaacctatth ttaagttgcc tcttggtatt aacattacaa	1980
atttttagagc cattottaca gccttttcac ctgctcaaga catttggggc acgtcagctg	2040
cagcctatth tgttggtat ttaaagccaa ctacattht gctcaagtat gatgaaaatg	2100
gtacaatcac agatgctgtt gattgttctc aaaatccact tgctgaactc aaatgctctg	2160
ttaagagctt tgagattgac aaaggaatth accagacctc taatttcagg gttgttccct	2220
caggagatgt tgtgagattc cctaataatta caaacttggtg tccttttgga gaggttttta	2280
atgctactaa attcccttct gtctatgcat gggagagaaa aaaaatttct aattgtgttg	2340
ctgattactc tgtgctctac aactcaacat ttttttcaac ctttaagtgc tatggcgtht	2400
ctgccactaa gttgaatgat ctttgcttct ccaatgtcta tgcagattct tttgtagtca	2460
agggagatga tgtaagacaa atagcgccag gacaaactgg tgttattgct gattataatt	2520
ataaattgcc agatgatttc atgggttggtg tccttgcttg gaatactagg aacattgatg	2580
ctacttcaac tggtaattat aattataaat ataggtatct tagacatggc aagcttaggc	2640
cctttgagag agacatatct aatgtgcctt tctccctga tggcaaacct tgcacccac	2700
ctgctcttaa ttgttattgg ccattaaatg attatggtth ttacaccact actggcattg	2760
gctaccaacc ttacagagtt gtagtactth cttttgaact tttaaatgca ccggccacgg	2820
tttggtggacc aaaattatcc actgacctta ttaagaacca gtgtgtcaat tttatttta	2880
atggactcac tggtagtgt gtgttaactc cttcttcaaa gagatttcaa ccatttcaac	2940
aatttggccg tgatgttht gatttactg attccgttcg agatcctaaa acatctgaaa	3000
tattagacat ttcacctgc tcttttgggg gtgtaagtgt aattacacct ggaacaaatg	3060
cttcatctga agttgctgtt ctatatcaag atgttaactg cactgatgtt tctacagcaa	3120
ttcatgcaga tcaactcaca ccagcttggc gcatatattc tactggaaac aatgtattcc	3180
agactcaagc aggctgtctt ataggagctg agcatgtcga cacttcttat gagtgcgaca	3240
ttcctattgg agctggcatt tgtgctagtt accatacagt ttctttatta cgtagtacta	3300
gccaaaaatc tattgtggct tatactatgt ctttaggtgc tgatagttca attgcttact	3360
ctaataacac cattgctata cctactaact tttcaattag cattactaca gaagtaatgc	3420
ctgtttctat ggctaaaacc tccgtagatt gtaatatgta catctgcgga gattctactg	3480
aatgtgctaa tttgcttctc caatatggta gcttttgac acaactaat cgtgcactct	3540
caggatttgc tgctgaacag gatcgcaaca cacgtgaagt gttcgctcaa gtcaaacaaa	3600
tgtacaaaac cccaactttg aaatattht gtggthttta tttttcacia atattacctg	3660
acctctaaa gccaaactaag aggtctthta ttgaggactt gctctthaat aagggtgacac	3720
tcgctgatgc tggcttcatg aagcaatatg gcgaatgcct aggtgatatt aatgctagag	3780
atctcatttg tgcgcagaag ttcaatggac ttacagtgtt gccacctctg ctactgatg	3840

atatgattgc tgcctacact gctgctctag ttagtggtac tgccactgct ggatggacat 3900
 ttggtgctgg cgctgctctt caaatacctt ttgctatgca aatggcatat aggttcaatg 3960
 gcattggagt tacccaaaat gttctctatg agaaccaaaa acaaatcgcc aaccaattta 4020
 acaaggcgat tagtcaaatt caagaatcac ttacaacaac atcaactgca ttgggcaagc 4080
 tgcaagacgt tgttaaccag aatgctcaag cattaaacac acttggttaa caacttagct 4140
 ctaattttgg tgcaatttca agtgtgctaa atgatatact ttgcgcactt gataaagtcg 4200
 aggcggaggt acaaattgac aggttaatta caggcagact tcaaagcctt caaacctatg 4260
 taacacaaca actaatcagg gctgctgaaa tcagggttc tgctaactct gctgctacta 4320
 aaatgtctga gtgtgttctt ggacaatcaa aaagagttga cttttgtgga aagggtacc 4380
 accttatgtc cttcccacaa gcagccccgc atggtgttgt cttcctacat gtcacgtatg 4440
 tgccatccca ggagaggaac ttcaccacag cgccagcaat ttgtcatgaa ggcaaagcat 4500
 acttccctcg tgaagggtt tttgtgttta atggcacttc ttggtttatt acacagagga 4560
 acttcttttc tccacaaata attactacag acaatacatt tgtctcagga aattgtgatg 4620
 tcgttattgg catcattaac aacacagttt atgatcctct gcaacctgag ctgcactcat 4680
 tcaaagaaga gctggacaag tacttcaaaa atcatacatc accagatgtt gatcttggcg 4740
 acatttcagg cattaacgct tctgtcgtca acattcaaaa agaaattgac cgcctcaatg 4800
 aggtcgctaa aaatttaaat gaatcactca ttgaccttca agaattggga aaatatgagc 4860
 aatatattaa atggcctcaa atactgtcaa tttattcaac agtggcgagt tccctagcac 4920
 tggcaatcat gatggctggt ctatctttat ggatgtgctc caatggatcg ttacaatgca 4980
 gaatttgcatt taaaggcct attttcttta gtttgaattt actgttattc ggtgtgcatt 5040
 tctatgtttg gtgagcgggt ttctgtgctc agagtgtgtt tttttatgt aatttaattt 5100
 ctttgtgagc tcctgtttag caggtcgtcc cttcagcaag gacacaaaaa gattttaatt 5160
 ttattaaaaa aaaaaaaaaa aaagaccggg aattcgatat caagcttatc gacctgcaga 5220
 tcgttcaaac atttggaat aaagtttctt aagattgaat cctgttgccg gtcttgcat 5280
 gattatcata taatttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat 5340
 gacgttattt atgagatggg tttttatgat tagagtcccg caattataca tttaatagc 5400
 gatagaaaac aaaatatagc gcgcaaacta ggataaatta tcgcgcgcgg tgtcatctat 5460
 gttactagat ctctagagtc tcaagcttgg cgcgcc 5496

<210> 31

<211> 1233

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của SARS gS- A/Indonesia/5/2005 H5 TM+CY

<400> 31

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
20 25 30

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
130 135 140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
195 200 205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225 230 235 240

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

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575

Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 675 680 685

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 690 695 700

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 705 710 715 720

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 725 730 735

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 740 745 750

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 755 760 765

Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 770 775 780

Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
 785 790 795 800

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 805 810 815

Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 820 825 830

Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 835 840 845

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
 850 855 860

Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
 865 870 875 880

Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
 885 890 895

Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
 900 905 910

Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
 915 920 925

Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
 930 935 940

Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
 945 950 955 960

Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
 965 970 975

Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
 980 985 990

Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
 995 1000 1005

Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp
1010 1015 1020

Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala
1025 1030 1035

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln
1040 1045 1050

Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys
1055 1060 1065

Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser
1070 1075 1080

Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr
1085 1090 1095

Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
1100 1105 1110

Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
1115 1120 1125

Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser
1130 1135 1140

Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val
1145 1150 1155

Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys
1160 1165 1170

Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
1175 1180 1185

Glu Gln Tyr Ile Lys Trp Pro Gln Ile Leu Ser Ile Tyr Ser Thr
1190 1195 1200

Val Ala Ser Ser Leu Ala Leu Ala Ile Met Met Ala Gly Leu Ser
1205 1210 1215

Leu Trp Met Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
1220 1225 1230

<210> 32

<211> 42

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> IF-RabG-S2+4.c

<400> 32

tctcagatct tcgccaaatt ccctatttac acgataccag ac

42

<210> 33

<211> 48

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> RabG+H5TM.r

<400> 33

ctgttgaata aattgacagt atttgatact tccccagtt cgggagac

48

<210> 34

<211> 1575

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Gen Rab G được tổng hợp(tương ứng với nt 3317-4891 từ số gia nhập ngân hàng gen

EF206707)

<400> 34

atggttcctc aggctctcct gtttgtaccc cttctggttt ttccattgtg ttttgggaaa 60

ttccctatatt acacgatacc agacaagctt ggtccctgga gcccgattga catacatcac 120

ctcagctgcc caaacaattt ggtagtggag gacgaaggat gcaccaacct gtcagggttc 180

tcctacatgg aacttaaagt tggatacatc ttagccataa aatgaacgg gttcacttgc 240

acaggcggtg tgacggaggc tgaaacctac actaacttcg ttggttatgt cacaaccacg 300

ttcaaaaagaa agcatttccg cccaacacca gatgcatgta gagccgcgta caactggaag 360

atggccggtg accccagata tgaagagtct ctacacaatc cgtaccctga ctaccgctgg 420

cttcgaactg taaaaaccac caaggagtct ctcgttatca tatctccaag tgtggcagat 480

ttggacccat atgacagatc ctttcaactg agggctcttc ctagcgggaa gtgctcagga 540

gtagcgggtg cttctaccta ctgctccact aaccacgatt acaccatttg gatgcccag 600

aatccgagac tagggatgtc ttgtgacatt ttaccaata gtagagggaa gagagcatcc 660

aaagggagtg agacttgcg ctttgtagat gaaagaggcc tatataagtc tttaaaagga 720

gcatgcaaac tcaagttatg tggagttcta ggacttagac ttatggatgg aacatgggtc 780

gcatgcaaaa catcaaatga aaccaaattg tgccctcccg atcagttggg gaacctgcac 840

gactttcgct cagacgaaat tgagcacctt gttgtagagg agttggtcag gaagagagag 900

gagtgtctgg atgcactaga gtccatcatg acaaccaagt cagtgagttt cagacgtctc 960

agtcatttaa gaaaacttgt ccctgggttt ggaaaagcat ataccatatt caacaagacc 1020
 ttgatggaag ccgatgctca ctacaagtca gtcagaactt ggaatgagat cctcccttca 1080
 aaagggtggt taagagttgg ggggaggtgt catcctcatg tgaacggggt gtttttcaat 1140
 ggtataatat taggacctga cggcaatgtc ttaatcccag agatgcaatc atccctcctc 1200
 cagcaacata tggagttggt ggaatcctcg gttatcccc ttgtgcaccc cctggcagac 1260
 ccgtctaccg ttttcaagga cggtgacgag gctgaggatt ttgttgaagt tcaccttccc 1320
 gatgtgcaca atcaggtctc aggagttgac ttgggtctcc cgaactgggg gaagtatgta 1380
 ttactgagtg caggggccct gactgccttg atgttgataa ttttcctgat gacatgttgt 1440
 agaagagtca atcgatcaga acctacgcaa cacaatctca gagggacagg gagggaggtg 1500
 tcagtcactc cccaaagcgg gaagatcata tcttcatggg aatcacacaa gagtgggggt 1560
 gagaccagac tgtga 1575

<210> 35
 <211> 29
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> IF-H5TM.c

<400> 35
 caaataactgt caatttattc aacagtggc 29

<210> 36
 <211> 4899
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Sản phẩm thiết kế 141

<400> 36
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60
 gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120
 aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180
 aatattcacc tactactggt ataaatcatt attaaacatt agagtaaaga aatatggatg 240
 ataagaacaa gagtagtgat attttgacaa caattttggt gcaacatttg agaaaatttt 300
 gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360
 aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaa atagttgtac 420
 aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480
 taaggatgac gcattagaga gatgtacat tagagaattt ttggcaagtc attaaaaaga 540
 aagaataaat tattttttaa attaaaagtt gagtcatttg attaaacatg tgattattta 600

atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt	660
taatttgaca tttgatcttt tcctatatat tgcccatag agtcagttaa ctcatTTTTA	720
tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaaacg	780
gtatatTTac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata	840
acatccaatc caaccaatca caacaatcct gatgagataa cccactTTaa gccacgcat	900
ctgtggcaca tctacattat ctaaatacaca cattcttcca cacatctgag ccacacaaaa	960
accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt	1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggaac gagctataca aggaaacgac gctagggaac aagctaacag tgaacgttgg	1140
gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg	1200
actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc	1260
aaggaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg	1320
gaagcttcac tgcacagagt ccttggtatct tggacgggag attcggTTaa ctatgcagca	1380
tctcgatttt tcggtttcga ccagatcgga tgtacctata gtattcggtt tcgaggagtt	1440
agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga	1560
tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt	1620
ctcctattta taatatggtt tgttattggt aattttgttc ttgtagaaga gcttaattaa	1680
tcgttgttgt tatgaaatac tatttgatg agatgaactg gtgtaatgta attcatttac	1740
ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg	1800
cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa	1860
gtggTTaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt	1920
tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatccct	1980
ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc	2040
ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg	2100
cgcgTTgga attactagcg cgtgtcgaca agcttgcatg ccggtcaaca tggTggagca	2160
cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat	2220
tgagactttt caacaaaggg taatatccgg aaacctctc ggattccatt gccagctat	2280
ctgtcacttt attgtgaaga tagtggaaaa ggaaggtggc tcctacaaat gccatcattg	2340
cgataaagga aaggccatcg ttgaagatgc ctctgcgcac agtggtccca aagatggacc	2400
ccccccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt	2460
ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga	2520

tacagtctca gaagaccaaa gggcaattga gactttttcaa caaagggtaa tatccggaaa	2580
cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga	2640
aggtggctcc tacaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc	2700
tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtag aaaaagaaga	2760
cgttccaacc acgtcttcaa agcaagtgga ttgatgtgat atctccactg acgtaaggga	2820
tgacgcacaa tcccactatc cttcgcaaga cccttcctct atataaggaa gttcatttca	2880
tttgagagag tattaaaaatc ttaatagggt ttgataaaaag cgaacgtggg gaaacccgaa	2940
ccaaaccttc ttctaaaactc tctctcatct ctcttaaagc aaacttctct cttgtctttc	3000
ttgctgagc gatcttcaac gttgtcagat cgtgcttcgg caccagtaca acgttttctt	3060
tcactgaagc gaaatcaaag atctctttgt ggacacgtag tgcggcgcca ttaaataacg	3120
tgtacttgct ctattcttgt cgggtgtggtc ttgggaaaag aaagcttgct ggaggctgct	3180
gttcagcccc atacattact tgttacgatt ctgctgaactc tcggcgggtg caatatctct	3240
acttctgctt gacgaggtat tggtgcctgt acttctttct tcttcttctt gctgattggt	3300
tctataagaa atctagtatt ttctttgaaa cagagttttc ccgtggtttt cgaacttgga	3360
gaaagattgt taagcttctg tatattctgc ccaaatttgt cgggcccattg gcgaaaaacg	3420
ttgcgatttt cggcttattg ttttctcttc ttgtgttggt tccttctcag atcttcgct	3480
gcaggctcct cagccaaaac gacacccccca tctgtctatc cactggcccc tggatctgct	3540
gccccaaacta actccatggt gaccctggga tgcttggtca agggctattt ccctgagcca	3600
gtgacagtga cctggaactc tggatccctg tccagcgggtg tgcacacctt ccagctgtc	3660
ctgcagtctg acctctacac tctgagcagc tcagtgaactg tcccctccag cacctggccc	3720
agcgagaccg tcacctgcaa cgttgccac ccggccagca gcaccaaggt ggacaagaaa	3780
attgtgcca gggattgtgg ttgtaagcct tgcataatgta cagtcccaga agtatcatct	3840
gtcttcatct tcccccaaa gcccaaggat gtgctacca ttactctgac tcctaaggctc	3900
acgtgtgttg tggtagacat cagcaaggat gatcccgagg tccagttcag ctggtttgta	3960
gatgatgtgg aggtgcacac agctcagacg caaccccggg aggagcagtt caacagcact	4020
ttccgctcag tcagtgaact tccatcatg caccaggact ggctcaatgg caaggagcga	4080
tcgctcacca tcaccatcac catcaccatc accattaaag gcctattttc tttagtttga	4140
atttactgtt attcgggtgtg catcttctatg tttggtgagc ggttttctgt gctcagagt	4200
tgtttatttt atgtaattta atttctttgt gagctcctgt ttagcaggctc gtcccttcag	4260
caaggacaca aaaagatttt aattttatta aaaaaaaaaa aaaaaaagac cgggaattcg	4320
atatcaagct tatcgacctg cagatcgctc aaacatttgg caataaagtt tcttaagatt	4380
gaatcctgtt gccggctctg cgatgattat catataattt ctgttgaatt acgttaagca	4440

tgtaataatt aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt 4500
 cccgcaatta tacatttaac acgcgataga aaacaaaata tagcgcgcaa actaggataa 4560
 attatcgcg cgggtgtcat ctatgttact agatctctag agtctcaagc ttggcgcgcc 4620
 cacgtgacta gtggcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt 4680
 tacccaactt aatcgccttg cagcacatcc ccctttcgcc agctggcgta atagcgaaga 4740
 ggcccgcaacc gatcgccctt cccaacagtt gcgcagcctg aatggcgaat gctagagcag 4800
 cttgagcttg gatcagattg tcgtttcccg ctttcagttt aaactatcag tgtttgacag 4860
 gatatatgg cggtgaaacc taagagaaaa gagcgttta 4899

<210> 37

<211> 3249

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Caset biểu hiện số 1074

<400> 37

gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60
 gaagacaaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga 120
 ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga aggtggctcc 180
 tacaaatgcc atcattgcga taaaggaaaag gccatcgttg aagatgcctc tgccgacagt 240
 ggtcccaaaag atggaccccc acccacgagg agcatcgttg aaaaagaaga cgttccaacc 300
 acgtcttcaa agcaagtga ttgatgtgat aacatgggtg agcacgacac acttgtctac 360
 tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420
 agggtaatat ccggaaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480
 aagatagtgg aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggcc 540
 atcgttgaag atgcctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc 600
 atcgtggaag aagaagacgt tccaaccacg tcttcaaagc aagtggattg atgtgatatc 660
 tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttctctata 720
 taaggaagtt cttttcattt ggagaggat taaaatctta ataggttttg ataaaagcga 780
 acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840
 cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900
 cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960
 ggcgccatta aataacgtgt acttgtccta ttcttgtcgg tgtggctcttg ggaaaagaaa 1020
 gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080
 gcgggtgcaa tatctctact tctgcttgac gaggtattgt tgctgtact tctttcttct 1140

tcttcttgct gattggttct ataagaaatc tagtatTTTT tttgaaacag agttttcccg	1200
tggttttcga acttgagaa agattgttaa gcttctgtat attctgccca aatttgtcgg	1260
gcccattggcg aaaaacgttg cgattttcgg cttattgttt tctcttcttg tgttggttcc	1320
ttctcagatc ttcgccaaat tccctattta cacgatacca gacaagcttg gtccctggag	1380
cccgattgac atacatcacc tcagctgccc aaacaatttg gtagtggagg acgaaggatg	1440
caccaacctg tcagggttct cctacatgga acttaaagtt ggatacatct tagccataaa	1500
aatgaacggg ttcacttgca caggcgttgt gacggaggct gaaacctaca ctaacttcgt	1560
tggttatgtc acaaccacgt tcaaaagaaa gcatttccgc ccaacaccag atgcatgtag	1620
agccgcgtac aactggaaga tggccggtga cccagatat gaagagtctc tacacaatcc	1680
gtaccctgac taccgctggc ttcgaactgt aaaaaccacc aaggagtctc tcgttatcat	1740
atctccaagt gtggcagatt tggaccata tgacagatcc cttcactcga gggcttccc	1800
tagcgggaag tgctcaggag tagcgggtgc ttctacctac tgctccacta accacgatta	1860
caccatttgg atgcccgaga atccgagact agggatgtct tgtgacattt ttaccaatag	1920
tagagggaag agagcatcca aagggtgta gacttgccgc tttgtagatg aaagaggcct	1980
atataagtct ttaaaaggag catgcaaact caagttatgt ggagttctag gacttagact	2040
tatggatgga acatgggtcg cgatgcaaac atcaaataa accaaatggg gccctccga	2100
tcagttggtg aacctgcacg actttcgctc agacgaaatt gagcaccttg ttgtagagga	2160
gttggtcagg aagagagagg agtgtctgga tgcactagag tccatcatga caaccaagtc	2220
agtgagtttc agacgtctca gtcatttaag aaaacttgct cctgggtttg gaaaagcata	2280
taccatattc aacaagacct tgatggaagc cgatgctcac tacaagtcag tcagaacttg	2340
gaatgagatc ctcccttcaa aagggtgttt aagagttggg gggaggtgtc atcctcatgt	2400
gaacggggtg tttttcaatg gtataatatt aggacctgac ggcaatgtct taatcccaga	2460
gatgcaatca tccctcctcc agcaacatat ggagttgttg gaatcctcgg ttatccccct	2520
tgtgcacccc ctggcagacc cgtctaccgt tttcaaggac ggtgacgagg ctgaggattt	2580
tgttgaagtt caccttcccg atgtgcacaa tcaggtctca ggagttgact tgggtctccc	2640
gaactggggg aagtatcaaa tactgtcaat ttattcaaca gtggcgagtt ccctagcact	2700
ggcaatcatg atggctggtc tatctttatg gatgtgctcc aatggatcgt tacaatgcag	2760
aatttgcatt taaaggccta ttttctttag tttgaattta ctgttattcg gtgtgcattt	2820
ctatgtttgg tgagcggttt tctgtgctca gagtgtgttt attttatgta atttaatttc	2880
tttgtgagct cctgttttagc aggtcgtccc ttcagcaagg acacaaaaag attttaattt	2940
tattaaaaaa aaaaaaaaaa aagaccggga attcgatatc aagcttatcg acctgcagat	3000
cgttcaaaca tttggcaata aagtttctta agattgaatc ctgttgccgg tcttgcatg	3060

attatcatat aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgcatg 3120
acgttatttta tgagatgggt ttttatgatt agagtcccg c aattatacat ttaatacgcg 3180
atagaaaaaca aaatatagcg cgcaaactag gataaattat cgcgcgcggt gtcatttatg 3240
ttactagat 3249

<210> 38

<211> 502

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của PDISP-Rab G- A/Indonesia/5/2005 H5 TM+CY

<400> 38

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 Lys Phe Pro Ile Tyr Thr Ile Pro
20 25 30

Asp Lys Leu Gly Pro Trp Ser Pro Ile Asp Ile His His Leu Ser Cys
35 40 45

Pro Asn Asn Leu Val Val Glu Asp Glu Gly Cys Thr Asn Leu Ser Gly
50 55 60

Phe Ser Tyr Met Glu Leu Lys Val Gly Tyr Ile Leu Ala Ile Lys Met
65 70 75 80

Asn Gly Phe Thr Cys Thr Gly Val Val Thr Glu Ala Glu Thr Tyr Thr
85 90 95

Asn Phe Val Gly Tyr Val Thr Thr Thr Phe Lys Arg Lys His Phe Arg
100 105 110

Pro Thr Pro Asp Ala Cys Arg Ala Ala Tyr Asn Trp Lys Met Ala Gly
115 120 125

Asp Pro Arg Tyr Glu Glu Ser Leu His Asn Pro Tyr Pro Asp Tyr Arg
130 135 140

Trp Leu Arg Thr Val Lys Thr Thr Lys Glu Ser Leu Val Ile Ile Ser
145 150 155 160

Pro Ser Val Ala Asp Leu Asp Pro Tyr Asp Arg Ser Leu His Ser Arg
165 170 175

Val Phe Pro Ser Gly Lys Cys Ser Gly Val Ala Val Ser Ser Thr Tyr

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

435

440

445

Asn Gln Val Ser Gly Val Asp Leu Gly Leu Pro Asn Trp Gly Lys Tyr
 450 455 460

Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala
 465 470 475 480

Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu
 485 490 495

Gln Cys Arg Ile Cys Ile
 500

<210> 39
 <211> 6863
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Sản phẩm thiết kế 144

<400> 39
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60
 gacgttttta atgtactgaa ttaacgccga atcccgggct ggtatattta tatgttgtca 120
 aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180
 aatattcacc tactactggt ataaatcatt attaaacatt agagtaaaga aatatggatg 240
 ataagaacaa gagtagtgat attttgacaa caattttggt gcaacatttg agaaaatttt 300
 gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360
 aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaa atagttgtac 420
 aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480
 taaggatgac gcattagaga gatgtacat tagagaattt ttggcaagtc attaaaaaga 540
 aagaataaat tatttttaaa attaaaagtt gagtcatttg attaaacatg tgattattta 600
 atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660
 taatttgaca tttgatcttt tcctatatat tgcccatag agtcagttaa ctcatTTTTA 720
 tatttcatag atcaaataag agaaataacg gtatattaat ccctccaaaa aaaaaaacg 780
 gtatattttac taaaaaatct aagccacgta ggaggataac aggatccccg taggaggata 840
 acatccaatc caaccaatca caacaatcct gatgagataa cccacttta gcccacgcat 900
 ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa 960
 accaatccac atcttttatca ccatttctat aaaaaatcac actttgtgag tctacacttt 1020
 gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag 1080

aaaatggaac gagctataca aggaaacgac gctaggggaac aagctaacag tgaacgttgg	1140
gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaag tccgagttgg	1200
actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc	1260
aaggaaaagct ggggtttcgg gaaagttgta tttaagagat atctcagata cgacaggacg	1320
gaagcttcac tgcacagagt ccttggatct tggacgggag attcggttaa ctatgcagca	1380
tctcgatttt tcggtttcga ccagatcggg tgtacctata gtattcgggtt tcgaggagtt	1440
agtatcacccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gcttgcccca atcgaagtgg aaagtaatgt atcaagagga	1560
tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt	1620
ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa	1680
tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac	1740
ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg	1800
cgtacaattg tcttatattt gaacaactaa aattgaacat cttttgccac aactttataa	1860
gtgggtaata tagctcaaata atatgggtcaa gttcaataga ttaataatgg aaatatcagt	1920
tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct	1980
ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc	2040
ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg	2100
cgcgttggga attactagcg cgtgtcgaca cgcggtggcg gccctagcag aaggcatggt	2160
gttgtgactc cgagggggtg cctcaaactc tatcttataa ccggcggtgga ggcatggagg	2220
caagggcatt ttggtaattt aagtagttag tggaaaatga cgtcatttac ttaaagacga	2280
agtcttgca caaggggggc ccacgccga ttttaatat accggcggtg cccacactta	2340
tcgcgagtgc tttagcacga gcggtccaga tttaaagtag aaaagttccc gccactagg	2400
gttaaagggtg ttcacactat aaaagcatat acgatgtgat ggtatttgat aaagcgtata	2460
ttgtatcagg tatttccgtc ggatacgaat tattcgtaca agcttcttaa gccggtcaac	2520
atggtggagc acgacacact tgtctactcc aaaaatatca aagatacagt ctcagaagac	2580
caaagggcaa ttgagacttt tcaacaaagg gtaatatccg gaaacctcct cggattccat	2640
tgcccagcta tctgtcactt tattgtgaag atagtggaaa aggaagggtg ctctacaaa	2700
tgccatcatt gcgataaagg aaaggccatc gttgaagatg cctctgccga cagtgggtccc	2760
aaagatggac cccacccac gaggagcatc gtggaaaaag aagacgttcc aaccacgtct	2820
tcaaagcaag tggattgatg tgataacatg gtggagcacg acacacttgt ctactccaaa	2880
aatatcaaag atacagtctc agaagaccaa agggcaattg agacttttca acaaagggtg	2940
atatccggaa acctcctcgg attccattgc ccagctatct gtcactttat tgtgaagata	3000

gtggaaaagg aaggtggctc ctacaaatgc catcattgcg ataaaggaaa ggccatcggt	3060
gaagatgcct ctgccgacag tgggtcccaa gatggacccc caccacagag gagcatcggt	3120
gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga tatctccact	3180
gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc tatataagga	3240
agttcatttc atttgagag gtattaaaat cttaataggt ttgataaaa gcgaacgtgg	3300
ggaaacccga accaaacctt cttctaaact ctctctcatc tctcttaaag caaacttctc	3360
tcttgctttt cttgcgtgag cgatcttcaa cgttgtcaga tcgtgcttcg gcaccagtac	3420
aacgtttttt ttcactgaag cgaaatcaaa gatctctttg tggacacgta gtgcggcgcc	3480
attaaataac gtgtacttgt cctattcttg tcggtgtggg cttgggaaaa gaaagcttgc	3540
tggaggctgc tgttcagccc catacattac ttgttacgat tctgctgact ttcggcggt	3600
gcaatatctc tacttctgct tgacgaggta ttgttgctg tacttctttt ttcttcttct	3660
tgctgattgg ttctataaga aatctagtat tttctttgaa acagagtttt cccgtgggtt	3720
tcgaacttgg agaaagattg ttaagcttct gtatattctg cccaaatttg tcgggccc	3780
ggcgaaaaac gttgcgattt tcggcttatt gttttctctt cttgtgttg ttcttctca	3840
gatcttcgcc tgcaggctcc tcagccaaaa cgacaccccc atctgtctat cactggccc	3900
ctggatctgc tgccaaaact aactccatgg tgaccctggg atgcctggc aagggctatt	3960
tccctgagcc agtgacagt acctggaact ctggatccct gtccagcgg gtgcacacct	4020
tcccagctgt cctgcagtct gacctctaca ctctgagcag ctgagtgact gtcccctcca	4080
gcacctggcc cagcgagacc gtcacctgca acgttgcca cccggccagc agcaccaagg	4140
tggacaagaa aattgtgccc agggattgtg gttgtaagcc ttgcatatgt acagtccag	4200
aagtatcatc tgtcttcac ttcccccaa agcccaagga tgtgctcacc attactctga	4260
ctcctaaggt cacgtgtgt gtggtagaca tcagcaagga tgatcccgag gtccagttca	4320
gctggtttgt agatgatgt gaggtgcaca cagctcagac gcaaccccg gaggagcagt	4380
tcaacagcac tttccgctca gtcagtgaac ttccatcat gcaccaggac tggctcaatg	4440
gcaaggaagg cctattttct ttagtttgaa ttactgtta ttcggtgtgc atttctatgt	4500
ttggtgagcg gttttctgtg ctgagagtgt gtttatttta tgtaatttaa tttctttgtg	4560
agctcctgtt tagcaggtcg tcccttcagc aaggacacaa aaagatttta attttattaa	4620
aaaaaaaaa aaaaaagacc gggaattcga tatcaagctt atcgacctgc agatcgttca	4680
aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc	4740
atataatttc tgttgaatta cgtaagcat gtaataatta acatgtaatg catgacgtta	4800
tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa	4860
aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta	4920

gatctctaga gtctcaagct tggcgcgggg taccgagctc gaattccgag tgtacttcaa	4980
gtcagttgga aatcaataaa atgattattht tatgaatata ttctattgtg caagtagata	5040
gaaattacat atgttacata acacacgaaa taaacaaaaa aacacaatcc aaaacaaaca	5100
ccccaaacaa aataaacacta tatatatcct cgtatgagga gaggcacgtt cagtgactcg	5160
acgattcccc agcaaaaaaa gtctccccgt cacacatata gtgggtgacg caattatctt	5220
caaagtaatc cttctgttga cttgtcattg ataacatcca gtcttcgtca ggattgcaaa	5280
gaattataga agggatccca ccttttattht tcttctthtt tccatattta gggttgacag	5340
tgaaatcaga ctggcaacct attaatgtct tccacaatgg gacgaacttg aaggggatgt	5400
cgtcgatgat attataggthg gcgtgttcat cgtagtthgt gaagtcgatg gtccccgttc	5460
agtagttgtg tgcctcgaga cttctagccc aggtggtctt tccggtacga gttggtccgc	5520
agatgtagag gctggggtgt ctgacccag tcttccctc atcctggtta gatcgcccat	5580
ccactcaagg tcagattgtg cttgatcgta ggagacagga tgtatgaaag tgtaggcac	5640
gatgcttaca tgatataggt gcgtctctct ccagttgtgc agatcttcgt ggcagcggag	5700
atctgattct gtgaagggcg acacgtactg ctcaggttht ggaggaaata atthgttggc	5760
tgaatattcc agccattgaa gctthgttgc ccattcatga gggaattctt cthtgatcat	5820
gtcaagatac tctccttag acgttgcatg ctggataata gttcgccatc gtgcgtcaga	5880
thtgcgagga gagacctat gatctcgga atctcctctg gthttaatat ctccgtcctt	5940
tgatatgtaa tcaaggactt gthtagagth tctagctggc tggatattag ggtgattthc	6000
ttcaaaatcg aaaaaagaag gatccctaath acaaggthth ttatcaagct ggataagagc	6060
atgatagthg gtagtgccat cthgatgaag ctgagaagca acaccaagga agaaaataag	6120
aaaaggthgt agththctcc agagaaactg gaataaatca tctctthtag atgagcacth	6180
gggtaggta aggaaaacat atthtagatth gagthctgaag thctthgctag cagaaggcat	6240
gthgtthgtga ctccgagggg thgcctcaaa ctctatctta taaccggcgt ggaggcatgg	6300
aggcaagggc atththgtaa thtaagtagth tagthgaaaa tgacgtcatt tactthaaaga	6360
cgaagthctg cgacaagggg ggccacgcc gaaththaat attaccggcg tggccccacc	6420
thtatcgag tgctthtagca cgagcggthc agaththaaag tagaaaagth cccgccact	6480
agggtthaaag gththcacac tataaaagca tatacgatgt gatggtatth gatggagcgt	6540
atattgtatc aggtatththc gtcggatacg aattatthct acggccggcc actagthgga	6600
ctggccgtcg thttacaacg tctgactgg gaaaaccctg gcgttaccca acttaatcg	6660
cttgagcac atccccctth cgccagctgg cgtaatagcg aagaggcccg caccgatcg	6720
ccttcccaac agthgcgag cctgaatggc gaatgctaga gcagctthgag cthggatcag	6780
atthgtctth cccgccttca gththaaacta tcatgthth acaggatata thggcgggta	6840

aacctaagag aaaagagcgt tta

6863

<210> 40

<211> 5279

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Caset biểu hiện số 1094

<400> 40

```

ctagcagaag gcatgttggt gtgactccga ggggttgccct caaactctat cttataaccg      60
gcgtggaggc atggaggcaa gggcattttg gtaatttaag tagttagtgg aaaatgacgt      120
catttactta aagacgaagt cttgcgacaa ggggggcccc cgccgaattt taatattacc      180
ggcgtggccc caccttatcg cgagtgtttt agcacgagcg gtccagattt aaagtagaaa      240
agttccccgc cactaggggt aaaggtgttc acactataaa agcatatacg atgtgatggt      300
atttgataaa gcgtatattg tatcaggtat ttccgtcgga tacgaattat tcgtacaagc      360
ttcttaagcc ggtcaacatg gtggagcacg acacacttgt ctactccaaa aatatcaaag      420
atacagtctc agaagaccaa agggcaattg agacttttca acaaagggtg atatccggaa      480
acctcctcgg attccattgc ccagctatct gtcactttat tgtgaagata gtggaaaagg      540
aaggtggctc ctacaaatgc catcattgag ataaaggaaa ggccatcgtt gaagatgcct      600
ctgccgacag tggccccaaa gatggacccc caccacgag gagcatcgtg gaaaaagaag      660
acgttccaac cacgtcttca aagcaagtgg attgatgtga taacatgggtg gagcacgaca      720
cacttgctca ctccaaaaat atcaaagata cagtctcaga agaccaaagg gcaattgaga      780
cttttcaaca aagggttaata tccggaaacc tcctcggatt ccattgcccc gctatctgtc      840
actttattgt gaagatagtg gaaaaggaag gtggctccta caaatgccat cattgcgata      900
aaggaaaggc catcgttgaa gatgcctctg ccgacagtgg tcccaaagat ggacccccac      960
ccacgaggag catcgttgaa aaagaagacg ttccaaccac gtcttcaaag caagtggatt     1020
gatgtgatat ctccactgac gtaagggatg acgcacaatc ccactatcct tcgcaagacc     1080
cttcctctat ataaggaagt tcatttcatt tggagaggta ttaaaatctt aataggtttt     1140
gataaaagcg aacgtgggga aaccggaacc aaaccttctt ctaaactctc tctcatctct     1200
cttaaagcaa acttctctct tgtctttctt gcgtgagcga tcttcaacgt tgtcagatcg     1260
tgcttcggca ccagtacaac gttttctttc actgaagcga aatcaaagat ctctttgtgg     1320
acacgtagtg cggcgccatt aaataacgtg tacttgctct attcttgctg gtgtgggtctt     1380
gggaaaagaa agcttgctgg aggctgctgt tcagcccat acattacttg ttacgattct     1440
gctgactttc ggcggttgca atatctctac ttctgcttga cgaggtattg ttgcctgtac     1500
ttctttcttc ttcttcttgc tgattgggtc tataagaaat ctagtatttt ctttgaaaca     1560

```

gagttttccc gtggttttcg aacttggaga aagattgtta agcttctgta tattctgccc	1620
aaatttgtcg ggcccatggc gaaaaacggt gcgattttcg gcttattggt ttctcttctt	1680
gtgttggttc cttctcagat cttcgccaaa ttccctattht acacgatacc agacaagcgt	1740
ggtccctgga gcccgattga catacatcac ctcagctgcc caaacaattt ggtagtggag	1800
gacgaaggat gcaccaacct gtcagggttc tcctacatgg aacttaaagt tggatacatc	1860
ttagccataa aatgaacgg gttcacttgc acaggcgttg tgacggaggc tgaaacctac	1920
actaacttcg ttggttatgt cacaaccacg ttcaaaagaa agcatttccg cccaacacca	1980
gatgcatgta gagccgcgta caactggaag atggccgggtg accccagata tgaagagtct	2040
ctacacaatc cgtaccctga ctaccgctgg cttcgaactg taaaaaccac caaggagtct	2100
ctcgttatca tatctccaag tgtggcagat ttggacccat atgacagatc ccttcactcg	2160
agggtcttcc ctagcgggaa gtgctcagga gtagcgggtg cttctaccta ctgctccact	2220
aaccacgatt acaccatttg gatgcccag aatccgagac tagggatgtc ttgtgacatt	2280
ttaccaata gtagagggaa gagagcatcc aaagggagtg agacttgagg ctttgtagat	2340
gaaagaggcc tatataagtc tttaaaagga gcatgcaaac tcaagttatg tggagttcta	2400
ggacttagac ttatggatgg aacatgggtc gcgatgcaaa catcaaatga aaccaaattg	2460
tgccctcccg atcagttggt gaacctgcac gactttcgtc cagacgaaat tgagcacctt	2520
gtttagagag agttggtcag gaagagagag gagtgtctgg atgcactaga gtccatcatg	2580
acaaccaagt cagtgagttt cagacgtctc agtcatttaa gaaaacttgt ccctgggttt	2640
ggaaaagcat ataccatatt caacaagacc ttgatggaag ccgatgctca ctacaagtca	2700
gtcagaactt ggaatgagat cctcccttca aaaggggtgt taagagttgg ggggaggtgt	2760
catcctcatg tgaacggggt gtttttcaat ggtataatat taggacctga cggcaatgtc	2820
ttaatcccag agatgcaatc atccctctc cagcaacata tggagttggt ggaatcctcg	2880
gttatcccc ttgtgcaccc cctggcagac ccgtctaccg ttttcaagga cggtgacgag	2940
gctgaggatt ttgttgaagt tcaccttccc gatgtgcaca atcaggtctc aggagttgac	3000
ttgggtctcc cgaactgggg gaagtatcaa atactgtcaa tttattcaac agtggcgagt	3060
tccctagcac tggcaatcat gatggctggt ctatctttat ggatgtgctc caatggatcg	3120
ttacaatgca gaatttgcac ttaaaggcct attttcttta gtttgaattt actgttattc	3180
ggtgtgcatt tctatgtttg gtgagcgggt ttctgtgctc agagtgtgtt tattttatgt	3240
aatttaattt ctttgtgagc tcctgtttag caggctgtcc cttcagcaag gacacaaaaa	3300
gattttaatt ttattaaaaa aaaaaaaaaa aaagaccggg aattcgatat caagcttatac	3360
gacctgcaga tcgttcaaac atttggcaat aaagtttctt aagattgaat cctgttgccg	3420
gtcttgcgat gattatcata taatttctgt tgaattacgt taagcatgta ataattaaca	3480

tgtaatgcat gacgttat ttt atgagatggg tttttatgat tagagtcccg caattataca 3540
 ttttaatacgc gatagaaaac aaaatatagc ggcgaaacta ggataaatta tcgcgcgcgg 3600
 tgtcatctat gttactagat ctctagagtc tcaagcttgg cgcggggtac cgagctcgaa 3660
 ttccgagtgat acttcaagtc agttggaaaat caataaaatg attat ttttat gaatatat ttt 3720
 cattgtgcaa gtagatagaa attacatatg ttacataaca cacgaaataa acaaaaaaac 3780
 acaatccaaa acaaacaccc caaacaaaat aacactatat atatcctcgt atgaggagag 3840
 gcacgttcag tgactcgacg attcccgagc aaaaaaagtc tccccgtcac acatatagtg 3900
 ggtgacgcaa ttatcttcaa agtaatcctt ctgttgactt gtcattgata acatccagtc 3960
 ttcgtcagga ttgcaaagaa ttatagaagg gatcccacct tttat ttttct tcttttttcc 4020
 atatttaggg ttgacagtga aatcagactg gcaacctatt aattgcttcc acaatgggac 4080
 gaacttgaag gggatgtcgt cgatgatatt ataggtggcg tgttcacgt agttggtgaa 4140
 gtcgatggtc ccgttccagt agttgtgtcg ccgagactt ctagcccagg tggctcttcc 4200
 ggtacgagtt ggtccgcaga tgtagaggct ggggtgtctg accccagtcc tccctcatc 4260
 ctggttagat cggccatcca ctcaaggcca gattgtgctt gatcgtagga gacaggatgt 4320
 atgaaagtgt aggcacgat gcttacatga tataggtgcg tctctctcca gttgtgcaga 4380
 tcttcgtggc agcggagatc tgattctgtg aagggcgaca cgtactgctc aggttgtgga 4440
 ggaaataatt tgttggctga atattccagc cattgaagct ttgttgccca ttcattgagg 4500
 aattcttctt tgatcatgtc aagatactcc tccttagacg ttgcagtctg gataatagtt 4560
 cgccatcgtg cgtcagat ttt gcgaggagag accttatgat ctcgaaatc tcctctgggt 4620
 ttaatatctc cgtcctttga tatgtaatca aggacttgtt tagagtttct agctggctgg 4680
 atattagggg gatttccttc aaaatcgaaa aaagaaggat ccctaataca aggtttttta 4740
 tcaagctgga taagagcatg atagtgggta gtgccatctt gatgaagctc agaagcaaca 4800
 ccaaggaaga aaataagaaa aggtgtgagt ttctcccaga gaaactggaa taaatcatct 4860
 ctttgagatg agcacttggg gtaggtaagg aaaacatatt tagattggag tctgaagtct 4920
 ttgctagcag aaggcatgtt gttgtgactc cgaggggttg cctcaaactc tatcttataa 4980
 ccggcgtgga ggcattggagg caagggcatt ttggtaattt aagtagttag tggaaaatga 5040
 cgtcatttac ttaaagacga agtcttgcca caaggggggc ccacgccgaa ttttaatat 5100
 accggcgtgg cccacactta tcgcgagtg tttagcacga gcggtccaga tttaaagtag 5160
 aaaagttccc gccactagg gttaaagggtg ttcacactat aaaagcatat acgatgtgat 5220
 ggtatttgat ggagcgtata ttgtatcagg tatttccgtc ggatacgaat tattcgtac 5279

<210> 41
 <211> 38
 <212> PRT

<213> Trình tự nhân tạo

<220>

<223> H5 (A/Indonesia/05/2005) TM/CT

<400> 41

Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu Ala
1 5 10 15

Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser Leu
20 25 30

Gln Cys Arg Ile Cys Ile
35

<210> 42

<211> 38

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> H3 (A/Brisbane/10/2007) TM/CT

<400> 42

Asp Trp Ile Leu Trp Ile Ser Phe Ala Ile Ser Cys Phe Leu Leu Cys
1 5 10 15

Val Ala Leu Leu Gly Phe Ile Met Trp Ala Cys Gln Lys Gly Asn Ile
20 25 30

Arg Cys Asn Ile Cys Ile
35

<210> 43

<211> 45

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> IF-Opt_EboGP.s2+4c

<400> 43

tctcagatct tcgccatccc tctgggagtg atccacaatt cgact

45

<210> 44

<211> 48

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> H5iTMCT+Opt_EboGP.r

<400> 44

ttgaataaat tgacagtatc tggcgccatc cggtccacca gttatcgt

48

<210> 45
 <211> 2031
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Gen GP tổng hợp được tối ưu hóa(tương ứng với nt 6039-8069 từ Sổ gia nhập Ngân hàng gen AY354458 với trình tự gen kiểu đại).

<400> 45

```

atgggggtga ccggaatcct tcagctccct cgcgataggt tcaaacggac gtccttcttt      60
ctctgggtga tcatcctggt tcagcgcacc ttctcaatcc ctctgggagt gatccacaat      120
tcgactctcc aggtgagcga ggtcgacaaa ctctgtgtgcc gcgataagtt gtccagcacc      180
aatcagctca gaagtgttgg cctcaacctg gaggggaatg gggtcgccac agacgtgcca      240
tccgcgacca agcgtgtggg ctttcggtcg ggtgttcccc cttaaagtcgt gaattacgag      300
gccggggagt gggctgaaaa ctgctacaac ctggaaatca agaaaccgga cggcagcgag      360
tgtctgcccc cggccccaga cgggatcagg ggtttcccgga ggtgccggta cgtgcacaag      420
gtctcgggaa ccggcccgtg cgcggcgcat tttgccttcc ataaggaggg cgcattcttt      480
ctctacgata gactggcctc cacggtcatc tatcgcgga ccaccttcgc ggaggggggtg      540
gtggcattcc tcatcctgcc gcaggctaag aaagatttct tcagctccca ccctctcagg      600
gaaccagtca acgccacaga ggatccctcc tcgggatatt acagcaccac aattcggtag      660
caggccacag gcttcgggac caatgagact gactacctgt tcgaggtgga caacctaa      720
tacgtgcagc tcgaatcgcg gtttaccctg cagttccttt tgcaacttaa cgagacgatc      780
tataccagtg gtaaacggag caacacgacc gggaagctca tctggaaagt aaaccggag      840
atcgacacca cgattggaga gtgggctttc tgggaaacca agaaaaacct gaccgggaag      900
atcaggtcgg aggagttgag cttcactgcc gtctccaata gagcaaaaaa catcagcggc      960
cagagccctg cccggacttc cagcgacca ggtaccaaca ccacgacaga ggaccacaag      1020
attatggcct cggagaattc ctctgcaatg gtccaggtgc atagccaggg gcgcgaagct      1080
gccgtctcgc acctgacaac cttggcaacg atttccacca gccacaacc acccagcaca      1140
aagcccgggc ccgacaactc cacccataat accccgtgtc ataagctgga catttccgaa      1200
gctacgcagg tggagcagca ccaccggagg accgacaacg actcaacagc tagcgatacc      1260
cctcccgcga ccacggcagc gggcccaccc aaggctgaga acaccaatac gagcaagggc      1320
accgacctcc tggacccgcg gactaccacg agccccaga atcacagcga aacggcgggc      1380
aacaacaaca cacaccatca ggatactggg gaggagtcgg cctccagcgg aaaactgggc      1440
ctgatcacia acaccatagc cggggctgcc gggctgatca ctggaggccg gagggcacgc      1500
agagaggcca ttgtgaacgc ccagcccaa tgcaacccaa acctccacta ctggaccacg      1560
caggacgagg gggccgctat cggcctggcc tggatcccat actttggggc cgccgctgag      1620

```

ggcatatata cggagggcct catgcacaat caggacggac tgatctgcgg actccgccag 1680
 cttgccaaacg agaccactca ggcgttgacg ctgtttctgc gggccactac cgagctcagg 1740
 acgttcagca ttctgaatcg gaaggcaatc gatttcctac tccagcgggtg gggcgggaca 1800
 tgccacatcc tgggaccgga ttgttgcatc gagccccacg actggaccaa gaacattaca 1860
 gataaaatcg accagatcat tcatgatttc gtggacaaga cactgccgga ccagggggac 1920
 aacgataact ggtggaccgg atggcgccag tggatcccag ccgggattgg cgtgacaggt 1980
 gtcattatcg ccgtgatcgc cctgttttgc atttgcaaat tcgttttcta g 2031

<210> 46
 <211> 45
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Opt_EboGP+H5iTMCT.c

<400> 46
 ggatggcgcc agatactgtc aatttattca acagtggcga gttcc 45

<210> 47
 <211> 4897
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Sản phẩm thiết kế 1192

<400> 47
 tggcaggata tattgtggtg taaacaaatt gacgcttaga caacttaata acacattgcg 60
 gacgttttta atgtactgaa ttaacgccga atccggggct ggtatattta tatgttgtca 120
 aataactcaa aaaccataaa agtttaagtt agcaagtgtg tacattttta cttgaacaaa 180
 aatattcacc tactactggt ataaatcatt attaaacatt agagtaaaga aatatggatg 240
 ataagaacaa gagtagtgat attttgacaa caattttggt gcaacatttg agaaaatttt 300
 gttgttctct cttttcattg gtcaaaaaca atagagagag aaaaaggaag agggagaata 360
 aaaacataat gtgagtatga gagagaaagt tgtacaaaag ttgtaccaa atagttgtac 420
 aaatatcatt gaggaatttg acaaaagcta cacaaataag ggttaattgc tgtaaataaa 480
 taaggatgac gcattagaga gatgtacat tagagaattt ttggcaagtc attaaaaaga 540
 aagaataaat tattttttaa attaaaagtt gagtcatttg attaaacatg tgattattta 600
 atgaattgat gaaagagttg gattaaagtt gtattagtaa ttagaatttg gtgtcaaatt 660
 taatttgaca tttgatcttt tcctatatat tgcccatag agtcagttaa ctcatTTTT 720
 tatttcatag atcaaataag agaaataacg gtatattaat cctccaaaa aaaaaaacg 780
 gtatatTTT tacaaaaatct aagccacgta ggaggataac aggatcccc taggaggata 840

acatccaatc caaccaatca caacaatcct gatgagataa cccactttaa gcccacgcat	900
ctgtggcaca tctacattat ctaaatcaca cattcttcca cacatctgag ccacacaaaa	960
accaatccac atctttatca cccattctat aaaaaatcac actttgtgag tctacacttt	1020
gattcccttc aaacacatac aaagagaaga gactaattaa ttaattaatc atcttgagag	1080
aaaatggaac gagctataca aggaaacgac gctaggggaaac aagctaacag tgaacgttgg	1140
gatggaggat caggaggtac cacttctccc ttcaaacttc ctgacgaaaag tccgagttgg	1200
actgagtggc ggctacataa cgatgagacg aattcgaatc aagataatcc ccttggtttc	1260
aaggaaaagct ggggtttcgg gaaagttgta tttaaagagat atctcagata cgacaggacg	1320
gaagcttcac tgcacagagt ccttgatct tggacgggag attcggttaa ctatgcagca	1380
tctcgatttt tcggtttcga ccagatcggg tgtacctata gtattcgggtt tcgaggagtt	1440
agtatcaccg tttctggagg gtcgcgaact cttcagcatc tctgtgagat ggcaattcgg	1500
tctaagcaag aactgctaca gcttgcccca atcgaaagtgg aaagtaatgt atcaagagga	1560
tgccctgaag gtactcaaac cttcgaaaaa gaaagcgagt aagttaaaat gcttcttcgt	1620
ctcctattta taatatggtt tgttattgtt aattttgttc ttgtagaaga gcttaattaa	1680
tcgttgttgt tatgaaatac tatttgtatg agatgaactg gtgtaatgta attcatttac	1740
ataagtggag tcagaatcag aatgtttcct ccataactaa ctagacatga agacctgccg	1800
cgtacaattg tcttataatt gaacaactaa aattgaacat cttttgccac aactttataa	1860
gtggttaata tagctcaaat atatggtcaa gttcaataga ttaataatgg aaatatcagt	1920
tatcgaaatt cattaacaat caacttaacg ttattaacta ctaattttat atcatcccct	1980
ttgataaatg atagtacacc aattaggaag gagcatgctc gcctaggaga ttgtcgtttc	2040
ccgccttcag tttgcaagct gctctagccg tgtagccaat acgcaaaccg cctctccccg	2100
cgcgttggga attactagcg cgtgtcgaca agcttgcagc ccggtcaaca tgggtggagca	2160
cgacacactt gtctactcca aaaatatcaa agatacagtc tcagaagacc aaagggcaat	2220
tgagactttt caacaaaggg taatatccgg aaacctcctc ggattccatt gccagctat	2280
ctgtcacttt attgtgaaga tagtggaata ggaaggtggc tcctacaaat gccatcattg	2340
cgataaagga aaggccatcg ttgaagatgc ctctgccgac agtgggtcca aagatggacc	2400
cccaccacg aggagcatcg tggaaaaaga agacgttcca accacgtctt caaagcaagt	2460
ggattgatgt gataacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga	2520
tacagtctca gaagacaaaa gggcaattga gacttttcaa caaagggtaa tatccggaaa	2580
cctcctcgga ttccattgcc cagctatctg tcactttatt gtgaagatag tggaaaagga	2640
aggtgggtcc taaaaatgcc atcattgcga taaaggaaag gccatcgttg aagatgcctc	2700
tgccgacagt ggtcccaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga	2760

cgttccaacc	acgtcttcaa	agcaagtgga	ttgatgtgat	atctccactg	acgtaaggga	2820
tgacgcacaa	tcccactatc	cttcgcaaga	cccttcctct	atataaggaa	gttcattttca	2880
tttgagagag	tattaaaaatc	ttaataggtt	ttgataaaaag	cgaacgtggg	gaaacccgaa	2940
ccaaaccttc	ttctaaactc	tctctcatct	ctcttaaagc	aaacttctct	cttgtctttc	3000
ttgcgtgagc	gatcttcaac	gttgtcagat	cgtgcttcgg	caccagtaca	acgttttctt	3060
tcactgaagc	gaaatcaaag	atctctttgt	ggacacgtag	tgcggcgcca	ttaaataacg	3120
tgtacttgtc	ctattcttgt	cggtgtggtc	ttgggaaaag	aaagcttgct	ggaggctgct	3180
gttcagcccc	atacattact	tgttacgatt	ctgctgactt	tcggcgggtg	caatatctct	3240
acttctgctt	gacgaggtat	tgttgccctgt	acttctttct	tcttcttctt	gctgattggt	3300
tctataagaa	atctagtatt	ttctttgaaa	cagagttttc	ccgtggtttt	cgaacttgga	3360
gaaagattgt	taagcttctg	tatattctgc	ccaaatttgt	cgggcccatg	gcgaaaaacg	3420
ttgcgatttt	cggcttattg	ttttctcttc	ttgtgttggt	tccttctcag	atcttcgccg	3480
cggctcctca	gccaaaacga	cacccccatc	tgtctatcca	ctggcccctg	gatctgctgc	3540
ccaaactaac	tccatgggtga	ccctgggatg	cctgggtcaag	ggctatttcc	ctgagccagt	3600
gacagtgacc	tggaactctg	gatccctgtc	cagcgggtgtg	cacaccttcc	cagctgtcct	3660
gcagtctgac	ctctacactc	tgagcagctc	agtgactgtc	ccctccagca	cctggcccag	3720
cgagaccgtc	acctgcaacg	ttgcccaccc	ggccagcagc	accaaggtgg	acaagaaaat	3780
tgtgcccagg	gattgtgggt	gtaagccttg	catatgtaca	gtcccagaag	tatcatctgt	3840
cttcatcttc	cccccaaagc	ccaaggatgt	gctcaccatt	actctgactc	ctaaggtcac	3900
gtgtgttggtg	gtagacatca	gcaaggatga	tcccagagtc	cagttcagct	ggttttaga	3960
tgatgtggag	gtgcacacag	ctcagacgca	accccgggag	gagcagttca	acagcacttt	4020
ccgctcagtc	agtgaacttc	ccatcatgca	ccaggactgg	ctcaatggca	aggagcgatc	4080
gctcaccatc	accatcacca	tcaccatcac	cattaaaggc	ctattttctt	tagtttgaat	4140
ttactgttat	tcgggtgtgca	tttctatgtt	tggtgagcgg	ttttctgtgc	tcagagtgtg	4200
tttattttat	gtaatttaat	ttctttgtga	gctcctgttt	agcaggtcgt	cccttcagca	4260
aggacacaaa	aagattttta	ttttattaaa	aaaaaaaaaa	aaaaagaccg	ggaattcgat	4320
atcaagctta	tcgacctgca	gatcgttcaa	acatttgga	ataaagtttc	ttaagattga	4380
atcctgttgc	cggctcttgcg	atgattatca	tataatttct	gttgaattac	gttaagcatg	4440
taataattaa	catgtaatgc	atgacgttat	ttatgagatg	ggtttttatg	attagagtcc	4500
cgcaattata	cattttaatac	gcgatagaaa	acaaaatata	gcgcgcaaac	taggataaat	4560
tatcgcgcg	ggtgtcatct	atgttactag	atctctagag	tctcaagctt	ggcgcgccca	4620
cgtgactagt	ggcactggcc	gtcgtttttac	aacgtcgtga	ctgggaaaac	cctggcggtta	4680

cccaacttaa tcgccttgca gcacatcccc ctttcgccag ctggcgtaat agcgaagagg 4740
cccgcaccga tcgcccttcc caacagttgc gcagcctgaa tggcgaatgc tagagcagct 4800
tgagcttgga tcagattgtc gtttcccgcc ttcagtttaa actatcagtg tttgacagga 4860
tatattggcg ggtaaaccta agagaaaaga gcgttta 4897

<210> 48
<211> 3780
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Caset biểu hiện số 1366

<400> 48
gtcaacatgg tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca 60
gaagacaaaa 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 cgttccaacc 300
acgtcttcaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac acttgtctac 360
tccaaaaata tcaaagatac agtctcagaa gaccaaaggg caattgagac ttttcaacaa 420
agggtaatat ccggaacct cctcggattc cattgcccag ctatctgtca ctttattgtg 480
aagatagtgg 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 ggagaggtat taaaatctta ataggttttg ataaaagcga 780
acgtggggaa acccgaacca aaccttcttc taaactctct ctcatctctc ttaaagcaaa 840
cttctctctt gtctttcttg cgtgagcgat cttcaacgtt gtcagatcgt gcttcggcac 900
cagtacaacg ttttctttca ctgaagcgaa atcaaagatc tctttgtgga cacgtagtgc 960
ggcgccatta aataacgtgt acttgtccta ttcttgctcg tgtggctctg ggaaaagaaa 1020
gcttgctgga ggctgctgtt cagccccata cattacttgt tacgattctg ctgactttcg 1080
gcggggtgcaa tatctctact tctgcttgac gaggtattgt tgccgtgact tctttcttct 1140
tcttcttgct gattggttct ataagaaatc tagtatcttc tttgaaacag agttttcccg 1200
tggttttcga acttgagaaa agattgttaa gcttctgtat attctgcca aatttgctcg 1260
gcccatggcg aaaaacgttg cgattttcgg cttattgttt tctcttcttg tgttggttcc 1320
ttctcagatc ttcgccatcc ctctgggagt gatccacaat tcgactctcc aggtgagcga 1380

ggctcgacaaa ctctgtgtgcc gcgataagtt gtccagcacc aatcagctca gaagtgttgg 1440
 cctcaacctg gaggggaatg gggctcgccac agacgtgccca tccgcgacca agcgttgggg 1500
 ctttcggtcg ggtgttcccc cttaaagtcgt gaattacgag gccggggagt gggctgaaaa 1560
 ctgctacaac ctggaaatca agaaacccga cggcagcgag tgtctgcccc cggccccaga 1620
 cgggatcagg ggtttcccgga ggtgccggta cgtgcacaag gtctcgggaa ccggcccgtg 1680
 cgccggcgat tttgccttcc ataaggaggg cgcattcttt ctctacgata gactggcctc 1740
 cacggtcatc tatcgcgga ccaccttcgc ggaggggggtg gtggcattcc tcatcctgcc 1800
 gcaggctaag aaagatttct tcagctccca ccctctcagg gaaccagtca acgccacaga 1860
 ggatccctcc tcgggatatt acagcaccac aattcggtag caggccacag gcttcgggac 1920
 caatgagact gagtacctgt tcgagggtga caacctaaac tacgtgcagc tcgaatcgcg 1980
 gtttaccocg cagttccttt tgcaacttaa cgagacgac tataccagtg gtaaaccggag 2040
 caacacgacc gggaagctca tctggaaagt aaaccgggag atcgacacca cgattggaga 2100
 gtgggctttc tgggaaacca agaaaaacct gaccgggaag atcaggtcgg aggagttgag 2160
 cttcactgcc gtctccaata gagcaaaaaa catcagcggc cagagccctg cccggacttc 2220
 cagcgacca ggtaccaaca ccacgacaga ggaccacaag attatggcct cggagaattc 2280
 ctctgcaatg gtccaggtgc atagccaggg gcgcgaagct gccgtctcgc acctgacaac 2340
 cttggcaacg atttccacca gccacaacc acccacgaca aagcccgggc ccgacaactc 2400
 cacccataat accccggtct ataagctgga catttccgaa gctacgcagg tggagcagca 2460
 ccaccggagg accgacaacg actcaacagc tagcgatacc cctcccgcca ccacggcagc 2520
 gggcccaccc aaggctgaga acaccaatac gagcaagggc accgacctcc tggaccccg 2580
 gactaccacg agccccaga atcacagcga aacggcgggc aacaacaaca cacaccatca 2640
 ggatactggg gaggagtcgg cctccagcgg aaaactgggc ctgatcaca acaccatagc 2700
 cggggctgcc gggctgatca ctggaggccg gagggcacgc agagaggcca ttgtgaacgc 2760
 ccagcccaaa tgcaaccaa acctccacta ctggaccacg caggacgagg gggccgctat 2820
 cggcctggcc tggatcccat actttgggcc cgccgctgag ggcatataca cggagggcct 2880
 catgcacaat caggacggac tgatctgcgg actccgccag cttgccaacg agaccactca 2940
 ggcgttgag ctgtttctgc gggccactac cgagctcagg acgttcagca ttctgaatcg 3000
 gaaggcaatc gatttcttac tccagcgggtg gggcgggaca tgccacatcc tgggacccga 3060
 ttgttgcatc gagccccacg actggacca gaacattaca gataaaatcg accagatcat 3120
 tcatgatttc gtggacaaga cactgccgga ccagggggac aacgataact ggtggaccgg 3180
 atggcgccag atactgtcaa ttattcaac agtggcgagt tccctagcac tggcaatcat 3240
 gatggctggg ctatctttat ggatgtgctc caatggatcg ttacaatgca gaatttgcac 3300

ttaaaggcct attttcttta gtttgaattt actgttattc ggtgtgcatt tctatgtttg 3360
 gtgagcgggt ttctgtgctc agagtgtgtt tattttatgt aatttaattt ctttgtgagc 3420
 tcctgttttag caggtcgtcc cttcagcaag gacacaaaaa gattttaatt ttattaaaaa 3480
 aaaaaaaaaa aaagaccggg aattcgatat caagcttatc gacctgcaga tcgttcaaac 3540
 atttggcaat aaagtttctt aagattgaat cctgttgccg gtcttgcat gattatcata 3600
 taatttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat gacgttattt 3660
 atgagatggg tttttatgat tagagtcccg caattataca tttaatacgc gatagaaaac 3720
 aaaatatagc gcgcaaaacta ggataaatta tcgcgcgcgg tgcatctat gttactagat 3780

<210> 49

<211> 679

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của PDISP-GP từ Zaire 95 Ebolavirut-H5 TM+CY
từ A/Indonesia/5/2005

<400> 49

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 Ile Pro Leu Gly Val Ile His Asn
20 25 30

Ser Thr Leu Gln Val Ser Glu Val Asp Lys Leu Val Cys Arg Asp Lys
35 40 45

Leu Ser Ser Thr Asn Gln Leu Arg Ser Val Gly Leu Asn Leu Glu Gly
50 55 60

Asn Gly Val Ala Thr Asp Val Pro Ser Ala Thr Lys Arg Trp Gly Phe
65 70 75 80

Arg Ser Gly Val Pro Pro Lys Val Val Asn Tyr Glu Ala Gly Glu Trp
85 90 95

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

Cys Leu Pro Ala Ala Pro Asp Gly Ile Arg Gly Phe Pro Arg Cys Arg
115 120 125

Tyr Val His Lys Val Ser Gly Thr Gly Pro Cys Ala Gly Asp Phe Ala
130 135 140

Phe His Lys Glu Gly Ala Phe Phe Leu Tyr Asp Arg Leu Ala Ser Thr
145 150 155 160

Val Ile Tyr Arg Gly Thr Thr Phe Ala Glu Gly Val Val Ala Phe Leu
165 170 175

Ile Leu Pro Gln Ala Lys Lys Asp Phe Phe Ser Ser His Pro Leu Arg
180 185 190

Glu Pro Val Asn Ala Thr Glu Asp Pro Ser Ser Gly Tyr Tyr Ser Thr
195 200 205

Thr Ile Arg Tyr Gln Ala Thr Gly Phe Gly Thr Asn Glu Thr Glu Tyr
210 215 220

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

Thr Pro Gln Phe Leu Leu Gln Leu Asn Glu Thr Ile Tyr Thr Ser Gly
245 250 255

Lys Arg Ser Asn Thr Thr Gly Lys Leu Ile Trp Lys Val Asn Pro Glu
260 265 270

Ile Asp Thr Thr Ile Gly Glu Trp Ala Phe Trp Glu Thr Lys Lys Asn
275 280 285

Leu Thr Arg Lys Ile Arg Ser Glu Glu Leu Ser Phe Thr Ala Val Ser
290 295 300

Asn Arg Ala Lys Asn Ile Ser Gly Gln Ser Pro Ala Arg Thr Ser Ser
305 310 315 320

Asp Pro Gly Thr Asn Thr Thr Thr Glu Asp His Lys Ile Met Ala Ser
325 330 335

Glu Asn Ser Ser Ala Met Val Gln Val His Ser Gln Gly Arg Glu Ala
340 345 350

Ala Val Ser His Leu Thr Thr Leu Ala Thr Ile Ser Thr Ser Pro Gln
355 360 365

Pro Pro Thr Thr Lys Pro Gly Pro Asp Asn Ser Thr His Asn Thr Pro
370 375 380

Val Tyr Lys Leu Asp Ile Ser Glu Ala Thr Gln Val Glu Gln His His
385 390 395 400

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

Ala Ile Met Met Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser
660 665 670

Leu Gln Cys Arg Ile Cys Ile
675