



(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)



1-0030261

(51)⁷ A61P 31/12; C12N 15/861

(13) B

(21) 1-2016-04355

(22) 13/05/2015

(86) PCT/US2015/030533 13/05/2015

(87) WO2015/175639 19/11/2015

(30) 61/992,649 13/05/2014 US

(45) 25/12/2021 405

(43) 27/03/2017 348A

(73) THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA (US)

3160 Chestnut Street, Suite 200, Philadelphia, PA 19104, United States of America

(72) WILSON, James, M. (US); TRETIAKOVA, Anna (US).

(74) Công ty TNHH Sở hữu công nghiệp Sao Bắc Đẩu (SAO BAC DAU IP CO.,LTD)

(54) VIRUT LIÊN QUAN ĐẾN ADENO (AAV) TÁI TỔ HỢP, CHẾ PHẨM VÀ DƯỢC PHẨM CHỨA VIRUT LIÊN QUAN ĐẾN ADENO TÁI TỔ HỢP NÀY

(57) Sáng chế đề cập đến virut liên quan đến adeno tái tổ hợp (AAV) có vỏ protein của AAV và được bọc bên trong đó axit nucleic khác loại mà biểu hiện hai cấu trúc kháng thể chức năng trong tế bào. Sáng chế cũng đề cập đến các kháng thể chứa chuỗi nặng và chuỗi nhẹ từ kháng thể khác loại. Theo một phương án, các kháng thể được cùng biểu hiện từ vectơ chứa: cat-xet biểu hiện thứ nhất mà mã hóa ít nhất khung đọc mở thứ nhất (ORF) đối với globulin miễn dịch thứ nhất dưới sự điều khiển của các trình tự điều khiển điều hòa mà định hướng sự biểu hiện của nó; và cat-xet biểu hiện thứ hai mà chứa ORF thứ hai, liên kết, và ORF thứ ba dưới sự điều khiển của các trình tự điều khiển điều hòa mà định hướng sự biểu hiện của nó, trong đó ORF thứ hai và thứ ba dùng cho cấu trúc globulin miễn dịch thứ hai và thứ ba. Vectơ đồng biểu hiện hai cấu trúc kháng thể này theo một phương án là AAV, trong đó các đoạn lặp đầu cuối đảo ngược (ITR) ở đầu 5' và 3' nằm cạnh cat-xet biểu hiện và trình tự điều hòa.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến virus liên quan đến adeno (AAV) tái tổ hợp, được phẩm chứa AAV tái tổ hợp và chế phẩm chứa hai kháng thể chức năng có các tính đặc hiệu khác nhau.

Tình trạng kỹ thuật của sáng chế

Các kháng thể đơn dòng đã được chứng minh là liệu pháp chữa bệnh hữu hiệu cho bệnh ung thư và các bệnh khác. Liệu pháp kháng thể hiện nay thường liên quan đến các chế độ điều trị lâu dài và cho dùng lặp lại, mà liên quan đến nhiều bất lợi, như các mức huyết thanh không thống nhất và giới hạn khoảng thời gian có hiệu quả cho mỗi lần cho dùng dẫn đến việc cần cho dùng lại thường xuyên và chi phí cao. Việc sử dụng các kháng thể làm công cụ chẩn đoán và phương thức chữa bệnh được thấy là sử dụng ngày càng nhiều trong những năm gần đây. Kháng thể đơn dòng được phê chuẩn bởi FDA thứ nhất để điều trị ung thư, Rituxan® (Rituximab) được phê chuẩn năm 1997 để điều trị cho các bệnh nhân mắc u lympho không Hodgkin và ngay sau đó năm 1995, phê chuẩn Herceptin®, kháng thể đơn dòng người hóa để điều trị cho người bệnh mắc bệnh ung thư vú di căn. Nhiều liệu pháp trên cơ sở kháng thể mà ở nhiều giai đoạn phát triển lâm sàng khác nhau đang thể hiện sự hứa hẹn. Với thành công của nhiều liệu pháp kháng thể đơn dòng khác nhau, đã được đề xuất thể hệ tiếp theo của dược phẩm sinh học sẽ liên quan đến hỗn hợp, tức là, hỗn hợp, của các kháng thể.

Một hạn chế đối với việc ứng dụng lâm sàng phổ biến của công nghệ kháng thể là thường cần lượng lớn kháng thể để có hiệu quả chữa bệnh và giá thành sản xuất là đáng kể. Tế bào buồng trứng chuột đồng Trung Quốc (Chinese Hamster Ovarian - CHO), tế bào u tủy SP20 và NSO2 là các dòng tế bào động vật có vú

được sử dụng phổ biến nhất cho việc sản xuất ở quy mô thương mại các protein người được glycosyl hóa như kháng thể. Hiệu suất thu được từ việc sản xuất dòng tế bào của động vật có vú thường nằm trong khoảng 50-250 mg/L trong 5-7 ngày nuôi cấy trong thiết bị lên men theo mẻ hoặc 300-1000 mg/L trong 7-12 ngày trong thiết bị lên men theo mẻ được cấp.

Virut liên quan đến adeno (AAV) là vectơ mong muốn để phân phối gen chữa bệnh nhờ hồ sơ an toàn của nó và khả năng biểu hiện gen dài hạn trong cơ thể. Vectơ AAV tái tổ hợp (rAAV) đã được sử dụng trước đó để biểu hiện các kháng thể chuỗi đơn và chiều dài đầy đủ *in vivo*. Do khả năng đóng gói gen chuyển bị giới hạn của AAV, có thử thách kỹ thuật để có hệ được điều hòa chặt chẽ để biểu hiện các chuỗi nặng và chuỗi nhẹ của kháng thể sử dụng vectơ AAV đơn nhằm tạo ra kháng thể chiều dài đầy đủ.

Vẫn có nhu cầu trong lĩnh vực này đối với việc phân phối hai kháng thể trong chế phẩm đơn để sử dụng trong chữa bệnh.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất virut liên quan đến adeno tái tổ hợp (AAV) có vỏ protein mà bọc bên trong đó axit nucleic khác loại mà biểu hiện hai kháng thể chức năng trong tế bào. Theo một phương án, AAV tái tổ hợp chứa ORF mã hóa chuỗi nhẹ globulin miễn dịch, ORF thứ hai mã hóa chuỗi nặng globulin miễn dịch thứ nhất và ORF thứ ba mã hóa chuỗi nặng thứ hai, nhờ đó các cấu trúc kháng thể chức năng được biểu hiện có hai chuỗi nặng khác nhau có các tính đặc hiệu khác nhau mà có chung chuỗi nhẹ. Theo một phương án, hai kháng thể có các tính đặc hiệu khác nhau được đồng biểu hiện, với kháng thể đặc hiệu kép, thứ ba có các tính đặc hiệu của hai kháng thể đơn đặc hiệu.

Theo một phương án, rAAV gồm: đoạn lặp đầu cuối đảo ngược (ITR) của AAV ở đầu 5'; cat-xet biểu hiện thứ nhất mà mã hóa ít nhất khung đọc mở thứ nhất (ORF) đối với globulin miễn dịch thứ nhất dưới sự điều khiển của các trình tự điều khiển điều hòa mà định hướng sự biểu hiện của nó; cat-xet biểu hiện thứ

hai mà chứa ORF thứ hai, liên kết, và ORF thứ ba dưới sự điều khiển của các trình tự điều khiển điều hòa mà định hướng biểu hiện của nó, trong đó ORF thứ hai và thứ ba mã hóa cho cấu trúc globulin miễn dịch thứ hai và thứ ba; và ITR của AAV tại đầu 3'.

Sáng chế đề xuất được phẩm chứa AAV tái tổ hợp mà biểu hiện ít nhất hai cấu trúc kháng thể chức năng và chất mang được dụng. Theo một phương án, ít nhất hai kháng thể chức năng có các tính đặc hiệu khác nhau. Tùy ý là, cũng được đồng biểu hiện là kháng thể đặc hiệu kép.

Sáng chế đề xuất chế phẩm chứa ít nhất hai kháng thể chức năng có các tính đặc hiệu khác nhau, trong đó mỗi kháng thể trong số các kháng thể có chuỗi nhẹ giống nhau và chuỗi nặng khác nhau. Chuỗi nhẹ là từ nguồn khác với chuỗi nặng cho một hoặc cả hai kháng thể. Theo một phương án, hai kháng thể đơn đặc hiệu chức năng và kháng thể hai chức năng được biểu hiện. Theo một phương án, tỷ lệ của các kháng thể bằng khoảng 25: khoảng 50: khoảng 25, dime đồng nhất: đặc hiệu kép: dime đồng nhất.

Các khía cạnh và lợi ích khác nữa của sáng chế sẽ trở nên rõ ràng hơn từ phần mô tả chi tiết các phương án thực hiện sáng chế sau đây.

Mô tả vắn tắt các hình vẽ

Hình 1A là hình vẽ minh họa một tổ hợp ví dụ cho vector biểu hiện hai cấu trúc kháng thể đơn đặc hiệu chứa chuỗi nặng thứ nhất và thứ hai và chuỗi nhẹ, mà có thể là từ kháng thể khác loại với một hoặc cả hai kháng thể mà chuỗi nặng thứ nhất và thứ hai bắt nguồn từ đó, và kháng thể đặc hiệu kép, kháng thể thứ ba. Tổ hợp này sử dụng vùng tăng cường chung mà theo hai hướng và tách biệt cat-xet biểu hiện thứ nhất và cat-xet biểu hiện thứ hai. Ba khung đọc mở (ORF) được minh họa. L đề cập đến liên kết. pA1 đề cập đến polyA thứ nhất và pA2 đề cập đến polyA thứ hai. MP1 đề cập đến vùng khởi đầu tối thiểu thứ nhất và MP2 đề

cập đến vùng khởi đầu tối thiểu thứ hai. PolyA và MP có thể giống hoặc khác nhau đối với mỗi cat-xet biểu hiện.

Hình 1B là hình vẽ minh họa tổ hợp ví dụ tùy ý cho vector biểu hiện hai cấu trúc kháng thể chứa chuỗi nặng thứ nhất và thứ hai và chuỗi nhẹ, mà có thể là từ kháng thể khác loại với một hoặc cả hai kháng thể mà chuỗi nặng thứ nhất và thứ hai bắt đầu từ đó, và kháng thể đặc hiệu kép, kháng thể thứ ba. Tổ hợp này sử dụng polyA chung. E1 đề cập đến vùng tăng cường thứ nhất và E2 đề cập đến vùng tăng cường thứ hai. Chúng có thể là các vùng tăng cường giống hoặc khác nhau đối với mỗi trong số các cat-xet biểu hiện. Tương tự MP1 và MP2 có thể giống hoặc khác nhau.

Hình 2 là hình vẽ minh họa phân tử axit nucleic được mang bởi plasmit để đóng gói vào trong vỏ protein của AAV, mà được sử dụng để đồng biểu hiện chuỗi nặng kháng TSG 101, chuỗi nặng cúm FI6 và chuỗi nhẹ FI6. Các chuỗi của kháng thể này sử dụng đoạn dẫn khác loại từ intelokin 2 (IL2). Vùng tăng cường CMV người được sử dụng kết hợp với các vùng khởi đầu CMV. Cat-xet biểu hiện xistron kép chứa vị trí nhận diện furin và trình tự liên kết 2A tách biệt ORF chứa vùng VL và CL của FI6 từ ORF chứa chuỗi nặng FI6. PolyA dùng cho cat-xet biểu hiện ở bên phải là polyA thymidin kinaza cắt ngắn. PolyA dùng cho cat-xet biểu hiện ở bên phải là trình tự polyA tổng hợp.

Hình 3 minh họa khả năng liên kết của kháng thể FI6v3k2 được đồng biểu hiện với kháng thể C05 từ AAV8 tái tổ hợp được điều chế như được mô tả ở đây. Các kết quả chứng minh sự liên kết được mong đợi vào HA chiều dài đầy đủ và thân HA đặc thù của FI6 và liên kết vào HA và chỉ đầu HA (không thân) đặc thù của C05.

Hình 4A-4B minh họa khả năng liên kết của kháng thể FI6v3k2 được đồng biểu hiện với kháng thể 1A6 (kháng-TSG 101) từ AAV8 tái tổ hợp được điều chế như mô tả ở đây. Hình 4A là biểu đồ cột thể hiện sự liên kết với protein A bất tổng kháng thể đơn dòng trong hỗn hợp (đối chứng âm được thể hiện bởi cột bên

trái, hỗn hợp kháng thể bởi cột bên phải). Hình 4B là sơ đồ thể hiện sự liên kết với peptit TSG101 chỉ bắt MAB chứa 1A6 chuỗi nặng (dòng trên). Các dữ liệu này chứng minh rằng khi được đồng biểu hiện với kháng thể FI6v3k2, 1A6 được duy trì tính đặc hiệu liên kết của kháng thể mà các chuỗi nặng của nó bắt nguồn từ đó.

Hình 5 minh họa mức độ biểu hiện hệ thống ở chuột được cho dùng FI6 được đồng biểu hiện từ vectơ AAV với kháng thể thứ hai với các liều dùng bằng 1×10^{11} bản sao bộ gen (genome copy- GC) hoặc 1×10^{10} GC.

Hình 6A-6B minh họa việc đánh giá của AAV9.BiD.FI6v3_CR8033mAb được phân phối trong cơ (IM) ở 1×10^{11} GC đối với sự bảo vệ chống lại thử thách với chủng cúm PR8. Hình 6A là đồ thị đường thể hiện phần trăm thay đổi về trọng lượng. Hình tròn thể hiện cấu trúc AAV9 có vùng khởi đầu theo hai hướng biểu hiện kháng thể đơn dòng FI6v3 và CR8033 tổng hợp có cùng chuỗi nhẹ khác loại. Hình vuông thể hiện đối chứng dương, tức là, AAV9 biểu hiện kháng thể đơn tuýp FI6 mà cũng được phân phối ở 1×10^{11} GC, và hình tam giác thể hiện các động vật tự nhiên. Hình 6B thể hiện sự sống sót sau thử thách.

Hình 7A-7B minh họa việc đánh giá của AAV9.BiD.FI6v3_CR8033mAb được phân phối trong cơ (IM) ở 1×10^{11} GC để bảo vệ chống lại thử thách với chủng cúm B/Lee/40. Hình 7A là đồ thị đường thể hiện phần trăm thay đổi về trọng lượng. Hình tròn thể hiện cấu trúc AAV9 có vùng khởi đầu hai hướng biểu hiện các kháng thể đơn dòng FI6 và CR8033 tổng hợp có cùng chuỗi nhẹ khác loại. Hình vuông thể hiện đối chứng dương, tức là, AAV9 biểu hiện kháng thể đơn tuýp CR8033 cũng được phân phối ở 1×10^{11} GC, và hình tam giác thể hiện các động vật tự nhiên. Hình 7B thể hiện sự sống sót sau thử thách.

Hình 8A là biểu đồ thể hiện sự bảo vệ trong mô hình chuột sau khi dùng AAV mà biểu hiện cả kháng thể đơn dòng FI6v3 và TCN, như được biểu hiện theo trọng lượng của chuột qua các ngày. Đường trên (các hình thoi) thể hiện liều 25 microgram ($\mu\text{g/mL}$) và đường dưới thể hiện 0,4 $\mu\text{g/mL}$.

Hình 8B là biểu đồ thể hiện sự bảo vệ trong mô hình chuột sau khi dùng AAV mà biểu hiện cả kháng thể đơn dòng FI6v3 và IA6, như được biểu hiện theo trọng lượng của chuột qua các ngày. Đường trên (các hình thoi) thể hiện liều 263,2 microgram ($\mu\text{g/mL}$) và đường dưới thể hiện 36,5 $\mu\text{g/mL}$.

Hình 8C là biểu đồ thể hiện sự bảo vệ trong mô hình chuột sau khi cho dùng AAV mà biểu hiện cả kháng thể đơn dòng FI6v3 và CR8033, như được biểu hiện theo trọng lượng của chuột qua các ngày. Đường trên (các hình thoi) thể hiện liều 126,3 microgram ($\mu\text{g/mL}$) và đường dưới thể hiện 6,9 $\mu\text{g/mL}$.

Mô tả chi tiết các phương án thực hiện sáng chế

Vector được đề xuất ở đây mà phân phối ít nhất hai kháng thể chức năng bằng cách đồng biểu hiện hai chuỗi nặng khác nhau và chuỗi nhẹ đơn mà khi được biểu hiện trong tế bào tạo thành hai kháng thể chức năng có các tính đặc hiệu khác nhau, tức là, mà nhận diện các kháng nguyên khác nhau (hoặc phối tử). Kháng thể chức năng thứ ba còn có thể được biểu hiện và là kháng thể đặc hiệu kép, có chuỗi nặng của mỗi kháng thể trong số hai kháng thể đơn đặc hiệu. Thông thường, kháng thể thứ ba được biểu hiện ở mức thấp hơn hai kháng thể đơn đặc hiệu. Vector có thể được sử dụng *in vivo* để sản xuất có hiệu quả các chế phẩm mà sẽ sử dụng ít nhất hai kháng thể hoặc tế bào chủ sản sinh kháng thể có thể được thiết kế để chứa cat-xet biểu hiện cho hai chuỗi nặng khác nhau và một loại chuỗi nhẹ. Do đó, sáng chế còn bao gồm tế bào chủ biểu hiện hỗn hợp của hai kháng thể đơn đặc hiệu, trong đó mỗi kháng thể có tính đặc hiệu khác biệt nhưng chứa cùng chuỗi nhẹ, và kháng thể thứ ba là kháng thể đặc hiệu kép. Theo một phương án mong muốn, vector được thiết kế để phân phối ba cấu trúc kháng thể khác nhau ở đối tượng mà được cho dùng vector.

Theo một phương án, vector là AAV tái tổ hợp mà được đóng gói trong vỏ protein của AAV mà phân tử axit nucleic chứa các trình tự mã hóa hai chuỗi nặng khác nhau và chuỗi nhẹ đơn, mà khi được đồng biểu hiện tạo thành hai kháng thể

đơn đặc hiệu chức năng, *tức là*, kháng thể thứ nhất với chuỗi nặng thứ nhất và chuỗi nhẹ và kháng thể thứ hai với chuỗi nặng thứ hai và chuỗi nhẹ, và kháng thể thứ ba mà có một trong số mỗi trong số các chuỗi nặng và cùng chuỗi nhẹ để tạo thành kháng thể đặc hiệu kép.

“Kháng thể chức năng” có thể là kháng thể hoặc globulin miễn dịch mà liên kết vào đích được chọn (ví dụ, kháng nguyên trên tế bào ung thư hoặc mầm bệnh, như virus, vi khuẩn, hoặc ký sinh trùng) với ái lực liên kết vừa đủ để tạo ra kết quả sinh lý mong muốn, mà có thể bảo vệ (ví dụ, gây miễn dịch thụ động) hoặc chữa bệnh.

Vector AAV được đề xuất ở đây có thể chứa 1, 2, hoặc 3 khung đọc mở (open reading frame - ORF) đối với lên đến mười vùng globulin miễn dịch. Như được sử dụng ở đây, “vùng globulin miễn dịch” đề cập đến vùng của chuỗi nặng hoặc chuỗi nhẹ của kháng thể như được xác định theo kháng thể chiều dài đầy đủ, thông thường. Cụ thể hơn là, kháng thể chiều dài đầy đủ chứa polypeptit chuỗi nặng (H) mà chứa bốn vùng: một vùng biến đổi (VH) đầu cuối N và ba vùng hằng định (CH1, CH2 và CH3) đầu cuối C và polypeptit chuỗi nhẹ (L) mà chứa hai vùng: một vùng biến đổi (VL) đầu cuối N và một vùng hằng định (CL) đầu cuối C. Vùng Fc chứa hai vùng (CH2 - CH3). Vùng Fab có thể chứa một vùng hằng định và một vùng biến đổi cho mỗi chuỗi nặng và chuỗi nhẹ.

Trong vector AAV được mô tả ở đây, hai polypeptit chuỗi nặng chiều dài đầy đủ có thể được biểu hiện (mỗi cái 4 vùng) và polypeptit của chuỗi nhẹ (hai vùng). Theo một phương án mong muốn, hai polypeptit của chuỗi nặng có các tính đặc hiệu khác nhau, tức là, hướng đến các đích khác nhau. Do đó, các vector dùng được một mình hoặc trong tổ hợp, để biểu hiện hỗn hợp của các kháng thể.

Như được sử dụng ở đây, “tính đặc hiệu khác nhau” chỉ ra rằng các cấu trúc globulin miễn dịch được viện dẫn (ví dụ, kháng thể chiều dài đầy đủ, chuỗi nặng, hoặc cấu trúc khác có khả năng liên kết đích đặc hiệu) liên kết với vị trí đích khác nhau. Thích hợp là, trong cấu trúc kháng thể được biểu hiện kép, hai tính đặc hiệu

là không chồng lấn và/hoặc không can thiệp, và có thể tùy ý tăng cường lẫn nhau. Hai cấu trúc kháng thể (globulin miễn dịch) như được mô tả ở đây tạo ra tính đặc hiệu khác nhau bằng cách liên kết với vị trí đích khác nhau trên cùng mầm bệnh hoặc vị trí đích (ví dụ, virus protein hoặc khối u). Các kháng nguyên đích khác nhau này có thể là các chủng khác nhau của cùng loại virus (ví dụ, hai chủng cúm khác nhau), hoặc hai kháng nguyên khác nhau (ví dụ, kháng virus và chống ung thư, hai cấu trúc chống ung thư khác nhau, trong số những cái khác). Ví dụ, polypeptit của chuỗi nặng thứ nhất có thể kết hợp với chuỗi nhẹ để tạo thành cấu trúc kháng thể có tính đặc hiệu thứ nhất, polypeptit của chuỗi nặng thứ hai có thể kết hợp với chuỗi nhẹ để tạo thành cấu trúc kháng thể thứ hai có tính đặc hiệu thứ hai, và chuỗi nặng thứ nhất và thứ hai có thể kết hợp với chuỗi nhẹ để tạo thành kháng thể đặc hiệu kép. Các kháng thể có thể tùy ý đều nhắm đến các vị trí kháng nguyên khác nhau (các epitop) trên đích riêng (ví dụ, các vị trí đích khác nhau trên mầm bệnh vi rút, vi khuẩn, nấm hoặc ký sinh trùng được chọn) hoặc đến các đích khác nhau. Ví dụ, các chuỗi nặng từ hai kháng thể có thể được nhắm đến virus cúm, và có thể được đồng biểu hiện để tạo thành hai kháng thể đơn đặc hiệu (ví dụ, các chuỗi nặng từ các virus cúm FI6, CR8033 và CO5 có thể được chọn) và được biểu hiện với chuỗi nhẹ được chọn, và kháng thể đặc hiệu kép. Các ví dụ của kháng thể cúm thích hợp và cấu trúc kháng thể của mầm bệnh chống lây nhiễm qua đường không khí khác và phương pháp để phân phối chúng được mô tả trong, ví dụ, WO 2012/145572A1. Các kháng thể có thể cũng được nhắm đến các đích khác nhau (ví dụ, kháng thể kháng virus, bao gồm nhiễm virus mạn tính, nhiễm virus liên quan đến ung thư, hoặc protein bề mặt tế bào chống ung thư khác nhau hoặc các đích khác. Các ví dụ của các đích virus thích hợp bao gồm protein của hemagglutinin cúm hoặc các protein virus khác, virus gây suy giảm miễn dịch ở người (HIV), virus gây u nhú ở người (HPV), virus Epstein-Barr, virus herpes ở người, virus hợp bào hô hấp, trong số những cái khác. Do đó, sáng chế đặc biệt thích hợp để dùng trong chữa bệnh và phòng bệnh thụ động mà các tổ hợp của các kháng thể là mong muốn.

Thuật ngữ “globulin miễn dịch” được sử dụng ở đây để bao gồm các kháng thể, và các mảnh chức năng của chúng. Các kháng thể có thể tồn tại ở nhiều dạng khác nhau bao gồm, ví dụ, các kháng thể đa dòng, các kháng thể đơn dòng, kháng thể vùng đơn camelid hóa, các kháng thể nội bào (“các nội kháng thể”), kháng thể tái tổ hợp, kháng thể đa đặc hiệu (đặc hiệu kép), các mảnh kháng thể, như, Fv, Fab, F(ab)₂, F(ab)₃, Fab', Fab'-SH, F(ab')₂, các kháng thể mảnh biến đổi chuỗi nhẹ đơn (scFv), cụm đôi/bis-scFv, Fc, pFc', scFvFc (hoặc scFv-Fc), disulfua Fv (dsfv), các kháng thể đặc hiệu kép (bc-scFv) như kháng thể BiTE; kháng thể camelid, kháng thể lại nổi bề mặt, kháng thể người hóa, kháng thể của người đầy đủ, kháng thể vùng đơn (sdAb, còn gọi là NANOBODY®), kháng thể thể khảm, các kháng thể thể khảm chứa ít nhất một vùng hằng định của người, và tương tự. "Mảnh kháng thể" đề cập đến ít nhất một phần của vùng biến đổi của globulin miễn dịch mà gắn vào đích của chúng, ví dụ, tế bào u. Theo một phương án, globulin miễn dịch là IgG. Tuy nhiên, các tuýp khác của globulin miễn dịch có thể được chọn. Theo phương án khác, tuýp phụ IgG được chọn là IgG1. Tuy nhiên, các isotyp khác có thể được chọn. Ngoài ra, bất kỳ trong số các alen của IgG1 có thể được chọn.

Thuật ngữ “khác loại” khi được sử dụng liên quan đến protein hoặc axit nucleic chỉ ra rằng protein hoặc axit nucleic chứa hai hoặc nhiều trình tự hoặc các trình tự con mà không tìm thấy cùng mối quan hệ giống nhau trong tự nhiên. Ví dụ, axit nucleic thường được sản xuất bằng cách tái tổ hợp, có hai hoặc nhiều trình tự từ các gen không liên quan được sắp xếp để tạo ra axit nucleic chức năng mới. Ví dụ, theo một phương án, axit nucleic có vùng khởi đầu từ một gen được sắp xếp để định hướng sự biểu hiện của trình tự mã hóa từ gen khác. Do đó, liên quan đến trình tự mã hóa, vùng khởi đầu là khác loại. Thuật ngữ “chuỗi nhẹ khác loại” là chuỗi nhẹ chứa vùng biến đổi và/hoặc vùng hằng định từ kháng thể mà có tính đặc hiệu đích khác với tính đặc hiệu của chuỗi nặng.

Hai hoặc nhiều ORF được mang bởi phân tử axit nucleic được đóng gói trong vector có thể được biểu hiện từ hai cat-xet biểu hiện, một hoặc cả hai mà có thể là xistron kép. Vì cat-xet biểu hiện chứa các chuỗi nặng từ hai kháng thể khác nhau, nên mong muốn là đưa vào sự biến đổi trình tự giữa hai trình tự chuỗi nặng để tối thiểu hóa khả năng tái tổ hợp đồng nhất. Thông thường có biến đổi đủ giữa vùng biến đổi của hai kháng thể (VH-Ab1 và VH-Ab2). Tuy nhiên, mong muốn là đảm bảo có sự biến đổi trình tự mã hóa đủ giữa các vùng hằng định của kháng thể thứ nhất (Ab1) và kháng thể thứ hai (Ab2), tốt nhất là trong mỗi trong số các vùng CH1, CH2, và CH3. Ví dụ, theo một phương án, vùng hằng định của chuỗi nặng của kháng thể thứ nhất có thể có trình tự của nt 1 đến 705 như nêu ở SEQ ID NO: 1 (mà mã hóa axit amin 1 - 233 nêu ở SEQ ID NO:2) hoặc trình tự mà có độ tương đồng nằm trong khoảng từ 95% 99% tương đồng với nó mà không đưa vào bất kỳ các thay đổi nào về axit amin. Theo một phương án, các biến đổi về trình tự của các vùng này được đưa vào trong việc tạo thành các codon đồng nghĩa (tức là, các biến đổi của trình tự axit nucleic được đưa vào mà không có bất kỳ thay đổi nào ở mức axit amin). Ví dụ, chuỗi nặng thứ hai có thể có các vùng hằng định mà khác nhau ít nhất 15%, ít nhất khoảng 25%, ít nhất khoảng 35% (tức là, độ tương đồng nằm trong khoảng từ 65% đến 85%) trên CH1, CH2 và/hoặc CH3.

Ngay khi đích và globulin miễn dịch được chọn, các trình tự mã hóa cho globulin miễn dịch được chọn (ví dụ, (các) chuỗi nặng và/hoặc chuỗi nhẹ) có thể thu được và/hoặc được tổng hợp. Phương pháp giải trình tự axit nucleic (ví dụ, ARN và ADN) là đã biết đối với người có hiểu biết trung bình trong lĩnh vực này. Ngay khi trình tự của axit nucleic là đã biết, axit amin được suy ra và sau đó, có các chương trình máy tính trên cơ sở trang tin và bán trên thị trường, cũng như các công ty dịch vụ mà dịch lại trình tự axit amin thành trình tự mã hóa axit nucleic. Tham khảo, ví dụ, backtranseq bởi EMBOSS, <http://www.ebi.ac.uk/Tools/st/>; Gene Infinity (http://www.geneinfinity.org/sms/sms_backtranslation.html); ExPasy (<http://www.expasy.org/tools/>). Theo một phương án, ARN và/hoặc trình tự mã

hóa cADN được thiết kế để biểu hiện tối ưu trong tế bào của người. Các phương pháp tổng hợp axit nucleic là đã biết đối với người có hiểu biết trung bình trong lĩnh vực và có thể được sử dụng cho toàn bộ, hoặc các phần, của cấu trúc axit nucleic được mô tả ở đây.

Các vùng mã hóa tối ưu hóa bởi codon có thể được thiết kế bằng nhiều phương pháp khác nhau. Sự tối ưu hóa này có thể được thực hiện sử dụng các phương pháp mà có trực tuyến (ví dụ, GeneArt,), các phương pháp đã công bố, hoặc công ty mà cung cấp dịch vụ tối ưu hóa codon, ví dụ, như DNA2.0 (Menlo Park, CA). Một thuật toán tối ưu hóa codon được mô tả, ví dụ, trong công bố đơn sáng chế số WO 2015/012924, được kết hợp vào đây bằng cách viện dẫn. Cũng tham khảo, ví dụ, công bố đơn sáng chế Mỹ số 2014/0032186 và công bố đơn sáng chế Mỹ số 2006/0136184. Thích hợp là, toàn bộ chiều dài của khung đọc mở (ORF) cho mỗi sản phẩm được cải biến. Tuy nhiên, theo một số phương án, chỉ có đoạn của ORF có thể được thay thế. Bằng cách sử dụng một trong số các phương pháp này, người ta có thể áp dụng tần suất cho trình tự polypeptit được cho bất kỳ, và tạo ra đoạn axit nucleic của vùng mã hóa tối ưu hóa bởi codon mà mã hóa polypeptit.

Có một số lựa chọn để tiến hành các thay đổi thực tế cho các codon hoặc để tổng hợp các vùng mã hóa tối ưu hóa codon được thiết kế như mô tả ở đây. Các cải biến hoặc tổng hợp này có thể được thực hiện bằng cách sử dụng các thao tác sinh học phân tử thông thường và tiêu chuẩn đã biết đối với người có hiểu biết trung bình trong lĩnh vực. Theo một nghiên cứu, các dãy của cặp oligonucleotit bổ sung có chiều dài bằng 80-90 nucleotit trong mỗi cặp và mở rộng chiều dài của trình tự mong muốn được tổng hợp bằng các phương pháp tiêu chuẩn. Các cặp oligonucleotit này được tổng hợp sao cho khi ủ, chúng tạo thành các đoạn sợi kép có 80-90 cặp bazơ, chứa các đầu dính liền, ví dụ, mỗi oligonucleotit trong cặp được tổng hợp để mở rộng 3, 4, 5, 6, 7, 8, 9, 10, hoặc nhiều bazơ trên vùng mà việc bổ sung cho oligonucleotit còn lại trong cặp. Các đầu sợi đơn của mỗi cặp

oligonucleotit được thiết kế để ủ với đầu sợi đơn của cặp oligonucleotit kia. Các cặp oligonucleotit được ủ, và khoảng năm đến sáu trong số các đoạn sợi kép này sau đó được ủ cùng nhau thông qua các đầu sợi đơn dính liền, và sau đó cho chúng ghép với nhau và tách dòng thành vector tách dòng vi khuẩn tiêu chuẩn, ví dụ, vector TOPO® được cung cấp bởi Invitrogen Corporation, Carlsbad, Calif. Cấu trúc sau đó được xếp trình tự bằng các phương pháp tiêu chuẩn. Một vài trong các cấu trúc này chứa 5 đến 6 đoạn của 80 đến 90 đoạn cặp bazơ được ghép nối với nhau, tức là, các đoạn có khoảng 500 cặp bazơ, được tạo ra, sao cho toàn bộ trình tự mong muốn được thể hiện trong các dãy cấu trúc plasmit. Việc chèn các plasmit này sau đó được cắt bằng enzym giới hạn thích hợp và được ghép nối với nhau để tạo thành cấu trúc cuối cùng. Cấu trúc cuối cùng này sau đó được tách dòng thành vector tách dòng vi khuẩn tiêu chuẩn, và được giải trình tự. Các phương pháp bổ sung luôn là hiển nhiên đối với người có hiểu biết trung bình trong lĩnh vực. Ngoài ra, việc tổng hợp gen dễ dàng có sẵn trên thị trường.

Tùy ý, việc thế axit amin có thể được đưa vào vùng hằng định chuỗi nặng để làm tăng tính đa dạng trình tự giữa hai chuỗi nặng của kháng thể và/hoặc nhằm mục đích khác. Phương pháp và chương trình máy tính để tạo ra sự sắp thẳng hàng này là đã có và đã biết đối với người có hiểu biết trung bình trong lĩnh vực. Các thay thế có thể còn được viết là (axit amin được xác định bởi một mã ký tự)-vị trí #- (axit amin được xác định bởi một mã ký tự) nhờ đó axit amin thứ nhất được thế axit amin và axit amin thứ hai là axit amin thế ở vị trí cụ thể. Thuật ngữ "thế" và "thế của axit amin" và "thế axit amin" như được sử dụng ở đây đề cập đến việc thay thế một axit amin trong trình tự axit amin bằng một axit amin khác, trong đó axit amin khác là khác với axit amin được thay thế. Phương pháp thay thế axit amin là đã biết đối với người có hiểu biết trung bình trong lĩnh vực và bao gồm, nhưng không giới hạn ở, các đột biến của trình tự nucleotit mã hóa trình tự axit amin. Phương pháp thực hiện thế axit amin trong IgG được mô tả, ví dụ, cho WO 2013/046704, được kết hợp vào đây bằng cách viện dẫn đến việc thảo luận của nó về kỹ thuật biến đổi axit amin.

Thuật ngữ "thể axit amin" và các từ đồng nghĩa của nó được mô tả ở trên là nhằm bao hàm sự biến đổi trình tự axit amin bằng cách thay thế axit amin bằng axit amin khác, axit amin thế. Thế có thể là việc thế bảo tồn. Thuật ngữ bảo tồn, đề cập đến hai axit amin, nghĩa là axit amin có chung đặc điểm đã biết bởi một trong số các người có hiểu biết trung bình trong lĩnh vực. Thuật ngữ không bảo tồn, đề cập đến hai axit amin, nghĩa là axit amin mà có sự khác nhau về ít nhất một đặc điểm đã biết bởi một trong số người có hiểu biết trung bình trong lĩnh vực. Ví dụ, các đặc điểm này có thể bao gồm axit amin có các chuỗi bên không axit kỵ nước, axit amin có các chuỗi bên kỵ nước (mà có thể còn khác nhau là axit hoặc không axit), axit amin có các chuỗi bên kỵ nước béo, axit amin có các chuỗi bên kỵ nước thơm, axit amin có các chuỗi bên trung tính phân cực, axit amin có các chuỗi bên tích điện, axit amin có các chuỗi bên axit tích điện, và axit amin có các chuỗi bên bazơ tích điện. Cả axit amin có trong tự nhiên và không có trong tự nhiên là đã biết trong lĩnh vực và có thể được sử dụng làm axit amin thế theo các phương án. Do đó, thể axit amin bảo tồn có thể liên quan đến việc thay đổi axit amin thứ nhất có chuỗi bên kỵ nước với axit amin khác có chuỗi bên kỵ nước; ngược lại thể axit amin không bảo tồn có thể liên quan đến việc thay đổi axit amin thứ nhất có chuỗi bên kỵ nước axit với axit amin khác có chuỗi bên khác, ví dụ, chuỗi bên kỵ nước bazơ hoặc chuỗi bên ưa nước. Tuy nhiên, thể bảo tồn hoặc không bảo tồn khác nữa được xác định bởi một trong số người có hiểu biết trung bình trong lĩnh vực. Theo các phương án khác nữa, việc thế ở vị trí được cho sẽ là cho axit amin, hoặc một trong số nhóm axit amin, mà hiển nhiên đối với một trong số người có hiểu biết trung bình trong lĩnh vực để hoàn thành mục tiêu được xác định ở đây.

Để biểu hiện vùng globulin miễn dịch được chọn, phân tử axit nucleic có thể được thiết kế mà chứa các codon mà được chọn để biểu hiện tối ưu globulin miễn dịch polypeptit ở các loài động vật có vú được chọn, ví dụ, con người. Ngoài ra, phân tử axit nucleic có thể bao gồm trình tự đoạn dẫn khác loại đối với mỗi chuỗi nặng và chuỗi nhẹ của kháng thể được chọn, mà mã hóa peptit của đoạn dẫn

tính hiệu IL-2 kiểu đại hoặc đột biến được dung hợp ngược dòng của các polypeptit nặng và nhẹ bao gồm vùng hằng định và vùng biến đổi. Tuy nhiên, trình tự đoạn dẫn khác loại khác có thể được thế cho một hoặc cả hai peptit tín hiệu IL-2. Peptit của đoạn dẫn/tín hiệu có thể giống hoặc khác nhau đối với mỗi cấu trúc globulin miễn dịch của chuỗi nặng và chuỗi nhẹ. Chúng có thể là trình tự tín hiệu được tìm thấy một cách tự nhiên trong globulin miễn dịch (ví dụ, IgG), hoặc có thể là từ nguồn khác loại. Các nguồn khác loại này có thể là xytokin (ví dụ, IL-2, IL12, IL18, hoặc tương tự), insulin, albumin, β -glucuronidaza, kiềm proteaza hoặc peptit tín hiệu kích thích bài tiết fibronectin, trong số những cái khác.

Như được sử dụng ở đây, “cat-xet biểu hiện” đề cập đến trình tự axit nucleic mà chứa ít nhất khung đọc mở thứ nhất (ORF) và tùy ý ORF thứ hai. ORF có thể chứa hai, ba, hoặc bốn vùng kháng thể. Ví dụ, ORF có thể chứa chuỗi nặng chiều dài đầy đủ. Tùy ý, ORF có thể chứa một hoặc hai vùng kháng thể. Ví dụ, ORF có thể chứa vùng biến đổi chuỗi nặng và vùng hằng định chuỗi nặng đơn. Trong ví dụ khác, ORF có thể chứa vùng biến đổi chuỗi nhẹ và vùng hằng định chuỗi nhẹ. Do đó, cat-xet biểu hiện có thể được thiết kế là xistron kép, tức là, để chứa trình tự điều hòa mà định hướng biểu hiện của các ORF trên đó từ trình tự điều hòa chung. Trong trường hợp này, hai ORF thường được tách bởi liên kết. Các liên kết thích hợp, như vị trí liên kết ribozym bên trong (IRES) và/hoặc liên kết peptit tự phân cắt furin-2a (F2a), [tham khảo, ví dụ, Radcliffe và Mitrophanous, Gene Therapy (2004), 11, 1673-1674] là đã biết trong lĩnh vực. Thích hợp là, ORF được liên kết tùy ý với các trình tự điều khiển điều hòa mà định hướng biểu hiện trong tế bào đích. Các trình tự điều khiển điều hòa này có thể bao gồm polyA, vùng khởi đầu, và vùng tăng cường. Để tạo điều kiện cho việc đồng biểu hiện từ vector AAV, ít nhất một trong số vùng tăng cường và/hoặc trình tự polyA có thể chung cat-xet biểu hiện thứ nhất và thứ hai.

Theo một phương án, rAAV được đóng gói trong vỏ protein của virut AAV được chọn, phân tử axit nucleic bao gồm: 5' ITR, cat-xet biểu hiện thứ nhất, vùng tăng cường hai hướng, và cat-xet biểu hiện thứ hai, trong đó vùng tăng cường hai hướng tách cat-xet biểu hiện thứ nhất và thứ hai, và 3' ITR. Hình 1A được đề xuất ở đây là ví dụ của phương án này. Ví dụ, theo phương án này, vùng khởi đầu thứ nhất cho cat-xet biểu hiện thứ nhất nằm ở bên trái của vùng tăng cường hai hướng, sau đó là ít nhất khung đọc mở thứ nhất, và trình tự polyA, và vùng khởi đầu thứ hai. Ngoài ra, vùng khởi đầu thứ hai cho cat-xet biểu hiện thứ hai nằm ở bên phải của vùng tăng cường hai hướng, sau đó là ít nhất khung đọc mở thứ hai và polyA. Các vùng khởi đầu thứ nhất và thứ hai và các trình tự polyA thứ nhất và thứ hai có thể là giống hoặc khác nhau. Vùng khởi đầu tối thiểu và/hoặc polyA tối thiểu có thể được chọn để tiết kiệm không gian. Thông thường, theo phương án này, mỗi vùng khởi đầu hoặc là nằm liền kề (hoặc là bên trái hoặc bên phải (hoặc 5' hoặc 3')) trình tự vùng tăng cường và các trình tự polyA nằm liền kề với các ITR, với các ORF ở giữa chúng. Trong khi Hình 1A minh họa, trật tự của các ORF có thể được thay đổi, như có thể vùng globulin miễn dịch được mã hóa nhờ đó. Ví dụ, các trình tự hằng định và biến đổi của chuỗi nhẹ có thể nằm ở bên trái của vùng tăng cường và hai chuỗi nặng có thể được mã hóa bởi các ORF nằm ở bên phải của vùng tăng cường. Tùy ý, một trong số các chuỗi nặng có thể nằm ở bên trái của vùng tăng cường và các ORF ở bên phải của vùng tăng cường bởi mã hóa chuỗi nặng thứ hai và chuỗi nhẹ. Tùy ý, có thể có cấu hình đối lập, và cat-xet biểu hiện ở bên trái của vùng tăng cường có thể là xistron kép. Tùy ý, phụ thuộc vào vùng được mã hóa nào, cat-xet biểu hiện đều có thể là một xistron (ví dụ, mã hóa hai adhesin miễn dịch), hoặc đều là xistron kép (ví dụ, mã hóa hai FAB đầy đủ).

Theo phương án khác, rAAV được bọc trong vỏ protein của virut AAV được chọn, phân tử axit nucleic bao gồm: 5' ITR, cat-xet biểu hiện thứ nhất, polyA mà có chức năng hai hướng, và cat-xet biểu hiện thứ hai, khi polyA hai hướng tách và các chức năng cho cả cat-xet biểu hiện thứ nhất và thứ hai, và 3' ITR. Hình 1B được đề xuất ở đây là ví dụ của phương án này. Theo phương án này,

vùng tăng cường thứ nhất và vùng khởi đầu thứ nhất (hoặc kết hợp vùng tăng cường/vùng khởi đầu) nằm ở bên phải của 5' ITR, sau đó là (các) ORF và polyA hai hướng. Cat-xet biểu hiện thứ hai được tách từ cat-xet biểu hiện thứ nhất bởi polyA hai hướng và được dịch mã theo hướng đối lập. Trong cat-xet biểu hiện này, vùng tăng cường và vùng khởi đầu (hoặc kết hợp đoàn khởi đầu/vùng tăng cường) nằm liền kề với 3' ITR và (các) ORF liền kề với polyA hai hướng. Trong khi Hình 1B minh họa, trật tự của các ORF có thể được thay đổi, vì có thể vùng globulin miễn dịch được mã hóa nhờ đó. Ví dụ, các trình tự hằng định và biến đổi của chuỗi nhẹ có thể nằm ở bên trái của polyA và hai chuỗi nặng có thể được mã hóa bởi (các) ORF nằm ở bên phải của polyA. Tùy ý, một trong số các chuỗi nặng có thể nằm ở bên trái của polyA và ORFs ở bên phải của polyA mã hóa chuỗi nặng thứ hai và chuỗi nhẹ. Tùy ý, cũng có thể có cấu hình đối lập, và cat-xet biểu hiện ở bên trái của polyA có thể là xistron kép. Tùy ý, phụ thuộc vào vùng được mã hóa nào, cat-xet biểu hiện đều có thể là một xistron (ví dụ, mã hóa hai adhesin miễn dịch), hoặc đều có thể là xistron kép.

Tùy ý, cấu trúc biểu hiện được ví dụ trên các Hình 1A và 1B và được mô tả ở đây có thể được sử dụng để đồng biểu hiện các cấu trúc globulin miễn dịch khác. Ví dụ, hai adhesin miễn dịch (IA) có thể được biểu hiện từ hai cat-xet biểu hiện một xistron. Adhesin miễn dịch bao gồm hình dạng của kháng thể mà được biểu hiện là khung đọc mở đơn chứa đơn vị mảnh biến đổi chuỗi đơn (scFv) (tức là, VH liên kết với VL hoặc VL liên kết với VH) được dung hợp vào vùng Fc (CH2-CH3), (ví dụ, VH-VL-CH2-CH3 hoặc VL-VH-CH2-CH3). Theo cách khác, lên đến bốn scFv có thể được biểu hiện từ hai cat-xet biểu hiện xistron kép. Theo phương án tùy ý, IA có thể được đồng biểu hiện với kháng thể chiều dài đầy đủ. Theo phương án tùy ý, một FAB hoàn chỉnh có thể được đồng biểu hiện với kháng thể chiều dài đầy đủ hoặc hai FAB hoàn chỉnh có thể được đồng biểu hiện. Theo phương án khác nữa, các tổ hợp khác của kháng thể chiều dài đầy đủ, IA, hoặc mảnh FAB có thể được đồng biểu hiện.

Các trình tự điều khiển điều hòa thích hợp có thể được chọn và được thu từ nhiều nguồn khác nhau. Theo một phương án, vùng khởi đầu tối thiểu và/hoặc polyA tối thiểu có thể được sử dụng để bảo tồn kích thước.

Như được sử dụng ở đây, thuật ngữ “vùng khởi đầu tối thiểu” nghĩa là trình tự ADN ngắn bao gồm hộp TATA và các trình tự khác mà dùng để xác định vị trí khởi đầu dịch mã, mà các yếu tố điều hòa được thêm vào đó để điều khiển biểu hiện. Theo một phương án, vùng khởi đầu đề cập đến trình tự nucleotit mà bao gồm vùng khởi đầu tối thiểu cộng với các yếu tố điều hòa mà có khả năng điều khiển biểu hiện của trình tự mã hóa hoặc ARN chức năng. Loại trình tự vùng khởi đầu này bao gồm các yếu tố ngược dòng đầu gần và đầu xa hơn, các yếu tố đầu xa thường đề cập đến các vùng tăng cường. Theo một phương án, vùng khởi đầu tối thiểu là vùng khởi đầu tối thiểu Cytomegalovirut (CMV). Theo phương án khác, vùng khởi đầu tối thiểu thu được từ CMV của người (hCMV) như vùng khởi đầu tối thiểu thu được từ vùng khởi đầu ngay trước của hCMV (tham khảo, US 20140127749, và Gossen và Bujard (Proc. Natl. Acad. Sci. USA, 1992, 89: 5547-5551), mà được kết hợp vào đây bằng cách viện dẫn). Theo phương án khác, vùng khởi đầu tối thiểu thu được từ nguồn virus như, ví dụ: SV40 các vùng khởi đầu trước hoặc sau, vùng khởi đầu ngay trước của cytomegalovirut (CMV), hoặc vùng khởi đầu trước của Rous Sarcomavirut (RSV) ; hoặc từ vùng khởi đầu tế bào nhân chuẩn, ví dụ, vùng khởi đầu beta actin (Ng, Nuc. Acid Res. 17:601-615, 1989; Quitsche và cộng sự, J. Biol. Chem. 264:9539-9545, 1989), vùng khởi đầu GADPH (Alexander, M. C. và cộng sự, Proc. Nat. Acad. Sci. USA 85:5092-5096, 1988, Ercolani, L. và cộng sự, J. Biol. Chem. 263:15335-15341, 1988), vùng khởi đầu TK-1 (thymidin kinaza), vùng khởi đầu HSP (protein sốc nhiệt), vùng khởi đầu UbB hoặc UbC, PGK, vùng khởi đầu Efl-alpha hoặc bất kỳ vùng khởi đầu nhân chuẩn chứa hộp TATA (Công bố đơn Mỹ số 2014/0094392). Theo phương án khác, vùng khởi đầu tối thiểu bao gồm vùng khởi đầu nhỏ, như vùng khởi đầu nhỏ CLDN5 được mô tả trong công bố đơn Mỹ số 2014/0065666. Theo phương án khác, vùng khởi đầu tối thiểu là vùng khởi đầu Thymidin kinaza (TK). Theo

một phương án, vùng khởi đầu tối thiểu là đặc hiệu với mô, như một trong số vùng khởi đầu đặc hiệu với tế bào cơ. vùng khởi đầu TnISlow tối thiểu, vùng khởi đầu TnIFast tối thiểu hoặc vùng khởi đầu kinaza creatin của cơ (Công bố đơn Mỹ số 2012/0282695). Mỗi tài liệu trong các tài liệu này được kết hợp vào đây bằng cách viện dẫn.

Theo một phương án, tín hiệu polyadenyl hóa (poly(A)) là tín hiệu poly(A) tối thiểu, tức là, trình tự tối thiểu cần để polyadenyl hóa đủ. Theo một phương án, poly(A) tối thiểu là poly(A) tổng hợp, như loại được mô tả trong Levitt và cộng sự, *Genes Dev.*, 1989 Jul, 3(7):1019-25; và Xia và cộng sự, *Nat Biotechnol.* 2002 Oct; 20(10):1006-10. Epub 16 tháng chín năm 2002. Theo phương án khác, poly(A) được chiết xuất từ beta-globin poly(A) của thỏ. Theo một phương án, polyA đóng vai trò hai hướng (An và cộng sự, 2006, *PNAS*, 103(49): 18662–18667. Theo một phương án, poly(A) thu được từ trình tự tín hiệu poly A trước SV40. Mỗi tài liệu trong các tài liệu này được kết hợp ở đây bằng cách viện dẫn.

Như được mô tả ở đây, theo một phương án, vùng tăng cường đơn, hoặc vùng tăng cường giống nhau, có thể điều hòa phiên mã của nhiều gen khác loại trong cấu trúc plasmit. Các vùng tăng cường khác nhau thích hợp để dùng trong sáng chế là đã biết trong lĩnh vực và bao gồm, ví dụ, vùng tăng cường sớm CMV, vùng tăng cường Hoxc8, nPE1 và nPE2. Các vùng tăng cường bổ sung hữu dụng ở đây được mô tả trong Andersson và cộng sự, *Nature*, tháng ba năm 2014, 507(7493):455-61, được kết hợp ở đây bằng cách viện dẫn. Các yếu tố vùng tăng cường khác nữa có thể bao gồm, ví dụ, vùng tăng cường apolipoprotein, vùng tăng cường của cá ngựa vằn, yếu tố vùng tăng cường GFAP, và các vùng tăng cường của mô tả cụ thể như được mô tả trong WO 2013/1555222, yếu tố điều hòa sau phiên mã của virus sau viêm gan ở chuột chũi. Ngoài ra, hoặc theo cách khác, yếu tố khác, ví dụ, xytomegalovirus (HCMV) ở người - vùng khởi đầu PDGR ngay đầu (IE)-lai hoặc vùng khởi đầu khác – các yếu tố vùng tăng cường có thể được chọn. Để tăng cường biểu hiện các yếu tố khác có thể là các intron (như intron

promega hoặc intron ảo globin miễn dịch ở người -globulin thể khảm ở gà). Các vùng khởi đầu và vùng tăng cường khác hữu dụng ở đây có thể tìm thấy trong Mammalian Promoter/Enhancer Database được tìm thấy ở <http://promoter.cdb.riken.jp/>.

Các cấu trúc được mô tả ở đây có thể còn chứa trình tự điều khiển biểu hiện hoặc trình tự điều hòa khác như, ví dụ, bao gồm các trình tự khởi đầu, kết thúc, vùng khởi đầu và vùng tăng cường thích hợp; các tín hiệu xử lý ARN hiệu quả như các tín hiệu ghép nối và polyadenyl hóa (polyA); các trình tự mà ổn định ARN thông tin tế bào chất; các trình tự mà tăng cường hiệu quả dịch mã (tức là, trình tự tương đồng Kozak); các trình tự mà tăng cường khả năng ổn định protein; và khi cần, các trình tự mà tăng cường sự tiết của sản phẩm được mã hóa. Vùng khởi đầu có thể được chọn từ trong số những vùng khởi đầu cơ định, vùng khởi đầu đặc hiệu với mô, vùng khởi đầu đặc hiệu với tế bào, vùng khởi đầu đáp ứng các tín hiệu sinh lý học, hoặc vùng khởi đầu điều hòa được [tham khảo, ví dụ, WO 2011/126868 và WO 2013/049492].

Các trình tự điều hòa này được “liên kết hoạt động” với các trình tự gen cấu trúc globulin miễn dịch. Như được sử dụng ở đây, thuật ngữ “liên kết hoạt động” đề cập đến cả các trình tự điều hòa biểu hiện mà liên tục với gen mong muốn và các trình tự điều hòa biểu hiện mà hoạt động ở trans hoặc với khoảng cách để điều khiển gen quan tâm.

Các ví dụ của vùng khởi đầu cơ định thích hợp để kiểm soát biểu hiện của vùng kháng thể bao gồm, nhưng không giới hạn ở các vùng khởi đầu β -actin của gà (CB) hoặc beta actin từ các loài khác, vùng khởi đầu cytomegalovirus (CMV) ở người, các vùng khởi đầu sớm và muộn của virus 40 ở khỉ (SV40), vùng khởi đầu U6, các vùng khởi đầu metallothionein, vùng khởi đầu EF1 α , vùng khởi đầu ubiquitin, vùng khởi đầu hypoxanthin phosphoribosyl transferase (HPRT), vùng khởi đầu dihydrofolate reductase (DHFR) (Scharfmann và cộng sự, Proc. Natl. Acad. Sci. USA 88:4626-4630 (1991), vùng khởi đầu adenosine deaminase, vùng

khởi đầu phosphoglyxerol kinaza (PGK), vùng khởi đầu pyruvat kinaza, vùng khởi đầu phosphoglyxerol mutaza, vùng khởi đầu β -actin (Lai và cộng sự, Proc. Natl. Acad. Sci. USA 86: 10006-10010 (1989), UbB, UbC, đoạn lặp đoạn cuối dài (LTR) của Moloney Leukemivirut và các retrovirut khác, vùng khởi đầu thymidin kinaza của Herpes Simplex Virut và các vùng khởi đầu cơ định khác đã biết đối với người có hiệu biết trung bình trong lĩnh vực. Các ví dụ của các vùng khởi đầu đặc hiệu với mô hoặc tế bào thích hợp để dùng trong sáng chế bao gồm, nhưng không giới hạn ở, endothelin-I (ET -I) và Flt-I, mà đặc trưng cho các tế bào nội mô, FoxJ1 (mà nhắm đích các tế bào có mao).

Các vùng khởi đầu cảm ứng thích hợp để điều khiển biểu hiện của vùng kháng thể bao gồm các vùng khởi đầu đáp ứng các chất ngoại sinh (ví dụ, các chất dược lý) hoặc các tín hiệu sinh lý. Các yếu tố đáp ứng này bao gồm, nhưng không giới hạn ở yếu tố đáp ứng giảm oxy huyết (HRE) mà liên kết HIF-1 α và β , yếu tố đáp ứng kim loại-ion như được mô tả bởi Mayo và cộng sự (1982, Cell 29:99-108); Brinster và cộng sự (1982, Nature 296:39-42) và Searle và cộng sự (1985, Mol. Cell. Biol. 5:1480-1489); hoặc yếu tố đáp ứng sốc nhiệt như được mô tả bởi Nouer và cộng sự (trong: Heat Shock Response, ed. Nouer, L., CRC, Boca Raton, Fla., pp167-220, 1991)

Theo một phương án, việc biểu hiện của khung đọc mở được điều khiển bởi vùng khởi đầu điều hòa được mà cho phép kiểm soát chặt chẽ đối với việc phiên mã của ORF (gen), ví dụ, chất dược lý, hoặc các yếu tố phiên mã được hoạt hóa bởi chất dược lý hoặc theo phương án tùy ý, các tín hiệu sinh lý. Các ví dụ của các vùng khởi đầu điều hòa được mà là các phức yếu tố phiên mã phụ thuộc phối tử mà có thể được sử dụng bao gồm, không giới hạn, các thành viên của siêu họ thụ thể nhân được hoạt hóa bởi phối tử tương ứng của chúng (ví dụ, glucocorticoid, estrogen, progestin, retinoid, ecdyson, và các chất tương tự và các chất đa hình của chúng) và rTTA được hoạt hóa bởi tetracyclin. Các ví dụ của các hệ này, bao gồm, không giới hạn, ARGENT™ Transcriptional Technology

(ARIAD Pharmaceuticals, Cambridge, Mass.). Các ví dụ của các hệ vùng khởi đầu này được mô tả, ví dụ, trong WO 2012/145572, được kết hợp vào đây bằng cách viện dẫn.

Các vùng khởi đầu khác nữa có thể bao gồm, ví dụ, vùng tăng cường/vùng khởi đầu ngay trước của xytomegalovirut (CMV) ở người, vùng tăng cường/vùng khởi đầu sớm của SV40, vùng khởi đầu JC polymovirut, các vùng khởi đầu protein tính chất bazơ của myelin (MBP) hoặc protein dạng sợi có tính chất axit của tế bào thần kinh đệm (GFAP), vùng khởi đầu tiềm ẩn (LAP) của virut gây bệnh mụn rộp (HSV-1), vùng khởi đầu lặp đoạn cuối dài (LTR) của virut hợp bào hô hấp (RSV), vùng khởi đầu đặc hiệu thần kinh (NSE), vùng khởi đầu của các nhân tố phát triển thu được từ tiểu huyết cầu (PDGF), hSYN, vùng khởi đầu của hormone tập trung hồng cầu (MCH), CBA, vùng khởi đầu của protein dạng sợi có tính chất axit của tế bào thần kinh đệm (GFAP), vùng khởi đầu của metalloprotein ma trận (MPP), và vùng khởi đầu của beta-actin ở gà. Các vùng khởi đầu có thể giống hoặc khác nhau đối với mỗi cat-xet biểu hiện.

Để dùng trong sản xuất vector virut AAV (ví dụ, (r) AAV tái tổ hợp), cat-xet biểu hiện có thể được mang trên vector thích hợp, ví dụ, plasmid, mà được đưa đến tế bào chủ đóng gói. Các plasmid hữu dụng trong sáng chế có thể được thiết kế sao cho chúng thích hợp để sao chép và đóng gói trong tế bào nhân sơ, tế bào động vật có vú, hoặc cả hai. Các kỹ thuật chuyển nhiễm thích hợp và các tế bào chủ đóng gói là đã biết và/hoặc có thể đã được thiết kế bởi một trong số người có hiểu biết trung bình trong lĩnh vực.

Phương pháp tạo và phân lập AAV thích hợp để sử dụng làm vector là đã biết trong lĩnh vực. Thường tham khảo, ví dụ, Grieger&Samulski, 2005, "Adeno-associated virus as a gene therapy vector: Vector development, production and clinical applications," *Adv. Biochem. Engin/Biotechnol.* 99: 119-145; Buning và cộng sự, 2008, "Recent developments in adeno-associated virus vector technology," *J. Gene Med.* 10:717-733; và các tài liệu tham khảo được liệt kê

dưới đây, mỗi tài liệu được kết hợp ở đây bằng cách viện dẫn toàn bộ. Để đóng gói gen chuyên vào virion, ITR là các thành phần AAV duy nhất cần trong *cis* trong cùng cấu trúc với phân tử axit nucleic chứa cat-xet biểu hiện. Các gen cap và rep có thể được tạo ra trong *trans*.

Như được mô tả ở trên, thuật ngữ “khoảng” khi được sử dụng để thay đổi giá trị số nghĩa là sai số bằng $\pm 10\%$, trừ khi có quy định khác.

Như được sử dụng trong toàn bộ bản mô tả và yêu cầu bảo hộ, thuật ngữ “bao gồm” và “chứa” và các biến thể của nó bao gồm, “hàm chứa”, “bao gồm”, “có chứa” và “bao hàm”, trong số các biến thể khác, là bao gồm các thành phần, yếu tố, toàn bộ, bước khác và tương tự. Thuật ngữ “gồm có” hoặc “gồm” là không bao gồm các thành phần, yếu tố, toàn bộ, bước khác và tương tự.

Thuật ngữ "giống nhau" hoặc phần trăm "tương đồng," trong bối cảnh hai hoặc nhiều các trình tự axit nucleic hoặc polypeptit, đề cập đến hai hoặc nhiều trình tự hoặc các tiểu trình tự mà giống nhau hoặc có phần trăm đặc hiệu của các gốc axit amin hoặc nucleotit mà giống nhau (tức là, độ tương đồng khoảng 70%, tốt hơn là 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, hoặc độ tương đồng cao hơn trên vùng đặc hiệu (ví dụ, bất kỳ một trong số các ORF biến đổi được đề xuất ở đây khi được so sánh và sắp hàng để tương đồng tối đa trên cửa sổ so sánh hoặc vùng được thiết kế) như được đo bằng cách sử dụng thuật toán so sánh trình tự BLAST hoặc BLAST 2.0 với các thông số mặc định được mô tả dưới đây, hoặc bằng cách sắp hàng thủ công và kiểm tra bằng mắt (tham khảo, ví dụ, trang web NCBI hoặc tương tự). Theo ví dụ khác, các trình tự polynucleotit có thể được so sánh sử dụng Fasta, chương trình trong GCG Phiên bản 6.1. Fasta tạo ra sắp hàng và phần trăm tương đồng trình tự của các vùng của lập nhất giữa các trình tự yêu cầu và tìm kiếm. Ví dụ, phần trăm tương đồng trình tự giữa các trình tự axit nucleic có thể được xác định bằng cách sử dụng Fasta với các thông số mặc định (kích thước chữ bằng 6 và yếu tố NOPAM đối với ma trận điểm) như được đề xuất trong GCG Phiên bản 6.1, được kết hợp vào đây bằng

cách viện dẫn. Thông thường, các chương trình này được sử dụng ở cài đặt mặc định, mặc dù người có hiểu biết trung bình trong lĩnh vực có thể thay đổi cài đặt này nếu cần. Tùy ý, một trong số người có hiểu biết trung bình trong lĩnh vực có thể sử dụng thuật toán hoặc chương trình máy tính khác mà tạo ra ít nhất mức tương đồng hoặc thuật toán như được đề xuất bởi thuật toán và chương trình được viện dẫn. Định nghĩa này còn đề cập đến, hoặc có thể được áp dụng cho, việc bổ sung trình tự. Định nghĩa này còn bao gồm các trình tự mà có xóa và/hoặc thêm, cũng như những trình tự mà có thể. Như được mô tả dưới đây, thuật toán được ưu tiên có thể tính cho các khoảng trống và tương tự. Tốt hơn là, tương đồng tồn tại trên vùng mà dài ít nhất khoảng 25, 50, 75, 100, 150, 200 axit amin hoặc nucleotit, và thường trên vùng mà dài 225, 250, 300, 350, 400, 450, 500 axit amin hoặc nucleotit hoặc trên chiều dài đầy đủ của axit amin hoặc các trình tự axit nucleic.

Thông thường, khi sắp hàng được chuẩn bị trên cơ sở trình tự axit amin, việc sắp hàng chứa sự chèn và xóa mà cũng được xác định liên quan đến trình tự AAV tham chiếu và việc đánh số các gốc axit amin trên cơ sở phạm vi tham chiếu được đề xuất cho việc sắp hàng. Tuy nhiên, bất kỳ trình tự AAV được cho có thể có ít các gốc axit amin hơn phạm vi tham chiếu. Theo sáng chế, khi bàn trình tự bố mẹ, thuật ngữ “vị trí giống nhau” hoặc “vị trí tương ứng” đề cập đến axit amin nằm ở số gốc giống nhau trong mỗi trong số các trình tự, so với phạm vi viện dẫn cho các trình tự được sắp hàng. Tuy nhiên, khi bỏ sắp hàng ra, mỗi trong số các protein có thể có các axit amin này nằm ở các số gốc khác nhau. Việc sắp hàng được thực hiện bằng cách sử dụng bất kỳ trong số các chương trình sắp hàng nhiều trình tự được công bố hoặc bán trên thị trường. Các chương trình sắp hàng tình tự là sẵn có cho các trình tự axit amin, ví dụ, các chương trình “Clustal X”, “MAP”, “PIMA”, “MSA”, “BLOCKMAKER”, “MEME”, và “Match-Box”. Thông thường, bất kỳ trong số các chương trình này được sử dụng ở cài đặt mặc định, mặc dù một trong số người có hiểu biết trung bình trong lĩnh vực có thể thay đổi các cài đặt này nếu cần. Tùy ý, một trong số người có hiểu biết trung bình trong lĩnh vực có thể sử dụng thuật toán hoặc chương trình máy tính khác mà tạo ra ít

nhất mức tương đồng hoặc sắp hàng như được tạo ra bởi thuật toán và chương trình được viện dẫn. Tham khảo, ví dụ, J. D. Thomson và cộng sự, Nucl. Acids. Res., “A comprehensive comparison of multiple sequence alignments”, 27(13):2682-2690 (1999).

Theo một phương án, các cat-xet biểu hiện được mô tả ở đây được thiết kế thành yếu tố di truyền (ví dụ, plasmit con thoi) mà chuyển các trình tự cấu trúc globulin miễn dịch được mang trên đó vào tế bào chủ bao để tạo ra vectơ virus. Theo một phương án, yếu tố di truyền được chọn có thể được phân phối đến tế bào bao AAV bằng phương pháp thích hợp, bao gồm các kỹ thuật chuyển nhiễm, điện di, phân phối liposom, dung hợp màng tế bào, viên nén được bao ADN tốc lực cao, nhiễm virus và dung hợp thể nguyên sinh. Các tế bào đóng gói AAV ổn định cũng có thể được tạo ra. Tùy ý, cat-xet biểu hiện có thể được sử dụng để tạo ra vectơ virus khác với AAV, hoặc để tạo ra hỗn hợp của các kháng thể trong cơ thể. Phương pháp được dùng để tạo các cấu trúc này là đã biết đối với chuyên gia trong xử lý axit nucleic và bao gồm công nghệ gen, công nghệ tái tổ hợp, và các kỹ thuật tổng hợp. Tham khảo, ví dụ, Molecular Cloning: A Laboratory Manual, ed. Green and Sambrook, Cold Spring Harbor Press, Cold Spring Harbor, NY (2012).

Vectơ AAV

Vectơ AAV tái tổ hợp (phân tử virus AAV) có thể chứa, được bọc trong vỏ protein của virus AAV, phân tử axit nucleic chứa 5' AAV ITR, cat-xet biểu hiện được mô tả ở đây và ITR của AAV tại đầu 3'. Như được mô tả ở đây, cat-xet biểu hiện có thể chứa các yếu tố điều hòa cho (các) khung đọc mở trong mỗi cat-xet biểu hiện và phân tử axit nucleic có thể tùy ý chứa các chi yếu tố điều hòa bổ sung.

Vectơ AAV có thể chứa đoạn lặp đầu cuối đảo ngược (ITR) ở đầu 5' của AAV chiều dài đầy đủ và ITR ở đầu 3' chiều dài đầy đủ. Phiên bản rút gọn của 5' ITR, gọi là Δ ITR, đã được mô tả trong đó trình tự D và vị trí vị trí đầu tiêu tan

(trs) bị xóa. Chữ viết tắt “sc” nghĩa là tự bổ sung. “AAV tự bổ sung” đề cập đến cấu trúc trong đó vùng mã hóa được mang bởi trình tự axit nucleic tái tổ hợp của AAV được thiết kế để tạo thành khuôn ADN sợi kép nội phân tử. Khi lây nhiễm, thay vì đợi cho tổng hợp thông qua tế bào của sợi thứ hai, hai nửa bổ sung của scAAV sẽ liên kết để tạo thành một đơn vị ADN sợi kép (dsADN) mà sẵn sàng để sao chép và phiên mã ngay. Tham khảo, ví dụ, D M McCarty và cộng sự, “Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis”, *Gene Therapy*, (Tháng tám năm 2001), Tập 8, Số 16, Trang 1248-1254. Các AAV tự bổ sung được mô tả trong, ví dụ, các patent Mỹ số 6,596,535; 7,125,717; và 7,456,683, mỗi trong số đó được kết hợp ở đây bằng cách viện dẫn toàn bộ.

Khi AAV dạng giả được tạo ra, các ITR được chọn từ nguồn mà khác nguồn AAV của vỏ protein. Ví dụ, các ITR của AAV2 có thể được chọn để dùng với vỏ protein của virus AAV có hiệu quả đặc biệt cho thụ thể tế bào, mô đích hoặc đích virus được chọn. Theo một phương án, các trình tự ITR từ AAV2, hoặc phiên bản bị xóa của chúng (Δ ITR), được dùng cho tiện và để thúc đẩy chấp nhận điều hòa. Tuy nhiên, các ITR từ các nguồn AAV khác có thể được chọn. Khi nguồn của các ITR là từ AAV2 và vỏ protein của virus AAV là từ nguồn AAV khác, vector thu được có thể được gọi là dạng giả. Tuy nhiên, các nguồn khác của các ITR của AAV có thể được sử dụng.

Nhiều vỏ protein của virus AAV được mô tả. Phương pháp tạo ra các vector AAV được mô tả rộng rãi trong các tài liệu và tài liệu sáng chế, bao gồm, ví dụ, WO 2003/042397; WO 2005/033321, WO 2006/110689; US 7588772 B2. Nguồn vỏ protein của virus AAV có thể được chọn từ AAV mà nhắm đến mô mong muốn. Ví dụ, AAV thích hợp có thể bao gồm, ví dụ, AAV9 [US 7,906,111; US 2011-0236353-A1], rh10 [WO 2003/042397] và/hoặc hu37 [tham khảo, ví dụ, US 7,906,111; US 2011-0236353-A1]. Tuy nhiên, AAV khác, bao gồm, ví dụ, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, [patent Mỹ 7790449;

patent Mỹ 7282199] và các loại khác. Tuy nhiên, các nguồn khác vỏ protein của virut AAV và các yếu tố virut khác có thể được chọn, như các cấu trúc globulin miễn dịch khác và các yếu tố vector khác.

Vector virut AAV sợi đơn được đề xuất. Phương pháp để tạo ra và phân lập các vector virut AAV thích hợp để phân phối cho đối tượng là đã biết trong lĩnh vực. Tham khảo, ví dụ, patent Mỹ 7790449; patent Mỹ 7282199; WO 2003/042397; WO 2005/033321, WO 2006/110689; và US 7588772 B2]. Trong một hệ, dòng tế bào tạo ra được chuyển nhiễm tức thời với cấu trúc mã hóa gen chuyển nằm cạnh các ITR và (các) cấu trúc mã hóa rep và cap. Trong hệ thứ hai, dòng tế bào đóng gói mà cung cấp ổn định rep và cap được chuyển nhiễm tức thời với cấu trúc mã hóa gen chuyển nằm cạnh các ITR. Trong mỗi hệ này, virion AAV được tạo ra đáp ứng lại chuyển nhiễm với adenovirut hoặc herpesvirut trợ giúp, đòi hỏi việc tách các rAAV khỏi bị nhiễm tạp virut. Gần đây, các hệ đã được phát triển mà không yêu cầu lây nhiễm với virut hỗ trợ để thu hồi AAV – các chức năng hỗ trợ cần thiết (tức là, adenovirut E1, E2a, VA, và E4 hoặc herpesvirut UL5, UL8, UL52, và UL29, và herpesvirut polymeraza) cũng được cung cấp, *in trans*, bởi hệ này. Trong các hệ mới hơn, các chức năng hỗ trợ có thể được cấp bởi chuyển nhiễm tức thời của các tế bào với các cấu trúc mã hóa chức năng hỗ trợ cần thiết, hoặc các tế bào có thể được thiết kế để chứa ổn định các gen mã hóa các chức năng hỗ trợ, biểu hiện trong đó có thể được kiểm soát ở mức độ dịch mã hoặc sau dịch mã. Theo hệ khác nữa, gen chuyển nằm cạnh các ITR và gen rep/cap được đưa vào tế bào côn trùng bằng cách gây nhiễm với các vector trên cơ sở baculovirut. Để kiểm tra trên các hệ sản xuất này, thường tham khảo, ví dụ, Zhang và cộng sự, 2009, "Adenovirus-adenovirus hybrid for large-scale recombinant adeno-associated virus production," Human Gene Therapy 20:922-929, nội dung của mỗi tài liệu được kết hợp ở đây bằng cách viện dẫn toàn bộ. Phương pháp tạo ra và sử dụng chúng và các hệ tạo AAV khác cũng được mô tả trong các patent Mỹ sau đây, nội dung của tài liệu được kết hợp ở đây bằng cách viện dẫn toàn bộ: 5,139,941; 5,741,683; 6,057,152; 6,204,059;

6,268,213; 6,491,907; 6,660,514; 6,951,753; 7,094,604; 7,172,893; 7,201,898; 7,229,823; và 7,439,065.

Sử dụng và chế độ trị liệu

rAAV, tốt hơn là được tạo huyền phù trong chất mang tương hợp về sinh lý học, có thể được dùng cho bệnh nhân là người hoặc động vật có vú không phải người. Các chất mang này có thể đã được chọn bởi một trong số người có hiểu biết trung bình trong lĩnh vực theo chỉ dẫn mà virus chuyên được nhắm đến đó. Ví dụ, một chất mang thích hợp bao gồm nước muối, mà có thể được phối chế với nhiều dung dịch đệm khác nhau (ví dụ, nước muối đệm phosphat). Các chất mang ví dụ khác bao gồm nước muối vô trùng, lactoza, sucroza, maltoza, và nước. Việc chọn chất mang không phải là giới hạn của sáng chế. Tùy ý, chế phẩm theo sáng chế có thể chứa, ngoài rAAV và (các) chất mang, các thành phần được thông thường khác, như chất bảo quản, hoặc chất ổn định hóa học.

Phương pháp sử dụng các rAAV này, ví dụ, để gây miễn dịch thụ động được mô tả, ví dụ, trong WO 2012/145572. Phương pháp phân phối và sử dụng khác sẽ là hiển nhiên đối với một trong số người có hiểu biết trung bình trong lĩnh vực. Ví dụ, chế độ trị liệu như được mô tả ở đây có thể chứa, ngoài một hoặc nhiều kết hợp được mô tả ở đây, kết hợp khác với một hoặc nhiều thuốc sinh học, thuốc phân tử nhỏ, chất hóa học trị liệu, các chất tăng cường miễn dịch, bức xạ, phẫu thuật, và tương tự. Thuốc sinh học như được mô tả ở đây, là dựa trên peptit, polypeptit, protein, enzym, phân tử axit nucleic, vector (bao gồm các vector virus), hoặc tương tự.

Trong liệu pháp kết hợp, cấu trúc globulin miễn dịch phân phối nhờ AAV được mô tả ở đây được cho dùng trước, trong, hoặc sau khi bắt đầu trị liệu với chất khác, cũng như kết hợp bất kỳ của chúng, tức là, trước và trong, trước và sau, trong và sau, hoặc trước, trong và sau khi bắt đầu trị liệu. Ví dụ, AAV có thể được cho dùng trong khoảng từ 1 đến 30 ngày, tốt hơn là từ 3 đến 20 ngày, tốt hơn nữa là từ 5 đến 12 ngày trước khi bắt đầu trị liệu bức xạ. Theo phương án khác của

sáng chế, hóa học trị liệu được cho dùng đồng thời với hoặc, tốt hơn nữa là, sau liệu pháp globulin miễn dịch qua trung gian AAV (kháng thể). Theo các phương án khác nữa, chế phẩm theo sáng chế có thể được kết hợp với các chất sinh học khác, ví dụ, thuốc kháng thể đơn dòng tái tổ hợp, các thể tiếp hợp kháng thể - thuốc, hoặc tương tự. Ngoài ra, các kết hợp của các cấu trúc globulin miễn dịch phân phối nhờ AAV hiệu quả như được bộc lộ ở trên có thể được sử dụng trong các chế độ trị liệu này.

Phương pháp hoặc đường dùng thích hợp bất kỳ có thể được sử dụng để cho dùng chế phẩm chứa AAV như được mô tả ở đây, và tùy ý, cho dùng đồng thời các thuốc hoạt tính khác hoặc các trị liệu chung với các kháng thể trung gian AAV được mô tả ở đây. Các đường dùng bao gồm, ví dụ, dùng toàn thân, qua đường miệng, trong tĩnh mạch, trong màng bụng, dưới da, hoặc trong cơ.

Các đích cho các cấu trúc globulin miễn dịch được mô tả ở đây có thể được chọn từ nhiều mầm bệnh khác nhau, bao gồm, ví dụ, các tác nhân gây nhiễm vi khuẩn, virus, nấm và ký sinh. Các đích thích hợp có thể còn bao gồm các kháng nguyên ung thư hoặc liên quan đến ung thư hoặc các đích tương tự. Các đích khác nữa có thể bao gồm tình trạng bệnh tự miễn như viêm khớp mạn tính tăng dần (RA) hoặc đa xơ cứng (MS).

Các ví dụ về đích virus bao gồm virus cúm từ họ orthomyxoviridae, mà bao gồm: Cúm A, Cúm B, và Cúm C. Các loại virus tuýp A là các mầm bệnh ở người độc hại nhất. Các kiểu huyết thanh của cúm A mà liên quan đến đại dịch bao gồm, H1N1, mà gây ra cúm Tây Ban Nha vào năm 1918, và cúm Lợn năm 2009; H2N2, mà gây ra cúm Châu Á năm 1957; H3N2, mà gây ra cúm Hồng Kông năm 1968; H5N1, mà gây ra cúm Gia cầm năm 2004; H7N7; H1N2; H9N2; H7N2; H7N3; và H10N7.

Các kháng thể trung hòa rộng kháng cúm A đã được mô tả. Như được sử dụng ở đây, “kháng thể trung hòa rộng” đề cập đến kháng thể trung hòa mà trung hòa nhiều chủng từ nhiều tuýp phụ. Ví dụ, CR6261 [The Scripps Institute/

Crucell] được mô tả là kháng thể đơn dòng mà liên kết với khoảng rộng của virut cúm bao gồm "cúm Tây Ban Nha" 1918 (SC1918/H1) và với virut thuộc lớp H5N1 của cúm gia cầm mà lây từ gà sang người ở Việt Nam vào năm 2004 (Viet04/H5). CR6261 nhận diện vùng xoắn ốc được duy trì mức cao trong thân gần màng của hemagglutinin, protein chủ yếu trên bề mặt của virut cúm này. Kháng thể này được mô tả trong WO 2010/130636, được kết hợp vào đây bằng cách viện dẫn. Kháng thể trung hòa khác, F10 [XOMA Ltd] được mô tả là hữu dụng để kháng H1N1 và H5N1. [Sui và cộng sự, Nature Structural and Molecular Biology (Sui, và cộng sự 2009, 16(3):265-73)]. Các kháng thể kháng cúm khác, ví dụ, Fab28 và Fab49, có thể được chọn. Tham khảo, ví dụ, WO 2010/140114 và WO 2009/115972, mà được kết hợp bằng cách viện dẫn. Kháng thể khác nữa, như loại được mô tả trong WO 2010/010466, công bố đơn sáng chế Mỹ US/2011/076265, và WO 2008/156763, có thể được dễ dàng chọn.

Các virut gây bệnh đích khác bao gồm, arenavirut (bao gồm funin, machupo, và Lassa), filovirut (bao gồm Marburg và Ebola), hantavirut, picornoviridae (bao gồm rhinovirut, echovirut), coronavirut, paramyxovirut, morbillivirut, virut hợp bào hô hấp, togavirut, coxsackievirut, parvovirut B19, virut gây bệnh khó thở ở trẻ em, adenovirut, reovirut, variola (Variola major (Bệnh đậu mùa)) và Vaccinia (Cowpox) từ họ poxvirut, và varicella-zoster (pseudorabies).

Sốt xuất huyết do virut gây ra bởi các thành viên của họ arenavirut (sốt Lassa) (họ cũng liên quan đến viêm màng não-màng nhện tăng lympho bào (LCM)), filovirut (ebolvirut), và hantavirut (puremala). Thành viên của picornavirut (phân họ của rhinovirut), kết hợp với lạnh thông thường ở người. Họ coronavirut, mà bao gồm nhiều virut không ở người như virut viêm phế quản truyền nhiễm (gia cầm), virut viêm dạ dày ruột truyền nhiễm trên heo (lợn), virut viêm não tủy ngưng kết hồng cầu ở lợn (lợn), virut viêm phúc mạc truyền nhiễm ở mèo (mèo), coronavirut gây bệnh đường ruột ở mèo (mèo), coronavirut ở chó

(chó). Coronavirut gây bệnh hô hấp ở người, được coi là liên kết với bệnh viêm gan cảm lạnh thông thường, không phải A, B hoặc C, và triệu chứng hô hấp cấp tính nặng (SARS). Họ paramyxovirut bao gồm paravirut cúm tuýp 1, paravirut cúm tuýp 3, paravirut cúm tuýp 3 ở bò, rubulavirut (virut bệnh quai bị, paravirut cúm tuýp 2, paravirut cúm tuýp 4, virut bệnh Newcastle (gà), bệnh dịch tả trâu bò, morbillivirut, mà bao gồm bệnh sởi và bệnh sài sốt chó, và pneumovirut, mà bao gồm virut hợp bào hô hấp (RSV). Họ parvovirut bao gồm parvovirut ở mèo (bệnh viêm ruột ở mèo), panleucopeniavirut ở mèo, parvovirut ở chó, và parvovirut ở lợn. Họ adenovirut bao gồm virut (EX, AD7, ARD, O.B.) mà gây ra các bệnh hô hấp.

Cấu trúc kháng thể trung hòa kháng mầm bệnh vi khuẩn có thể còn được chọn để dùng theo sáng chế. Theo một phương án, cấu trúc kháng thể trung hòa được định hướng kháng chính vi khuẩn. Theo phương án khác, cấu trúc kháng thể trung hòa được định hướng kháng độc tố sinh ra bởi vi khuẩn. Các ví dụ về mầm bệnh vi khuẩn trong không khí bao gồm, ví dụ, *Neisseria meningitidis* (viêm màng não), *Klebsiella pneumonia* (viêm phổi), *Pseudomonas aeruginosa* (viêm phổi), *Pseudomonas pseudomallei* (viêm phổi), *Pseudomonas mallei* (viêm phổi), *Acinetobacter* (viêm phổi), *Moraxella catarrhalis*, *Moraxella lacunata*, *Alkaligenes*, *Cardiobacterium*, *Haemophilus influenzae* (cúm), *Haemophilusvirut* gây bệnh khó thở ở trẻ em, *Bordetella pertussis* (ho gà), *Francisella tularensis* (viêm phổi/sốt), *Legionella pneumonia* (bệnh nhiễm trùng phổi), *Chlamydia psittaci* (viêm phổi), *Chlamydia pneumoniae* (viêm phổi), *Mycobacterium tuberculosis* (bệnh lao (TB)), *Mycobacterium kansasii* (TB), *Mycobacterium avium* (viêm phổi), *Nocardia asteroides* (viêm phổi), *Bacillus anthracis* (bệnh than), *Staphylococcus aureus* (viêm phổi), *Streptococcus pyogenes* (bệnh ban đỏ), *Streptococcus pneumoniae* (viêm phổi), *Corynevi khuẩn bệnh bạch cầu* (bệnh bạch cầu), *Mycoplasma pneumoniae* (viêm phổi).

Các tác nhân gây ra bệnh than là độc tố được tạo ra bởi *Bacillus anthracis*. Các kháng thể trung hòa kháng chất bảo vệ (PA), là một trong ba peptit mà tạo thành biến độc tố, được mô tả. Hai polypeptit khác bao gồm yếu tố gây chết (LF) và yếu tố gây phù (EF). Các kháng thể trung hòa kháng PA được mô tả là hiệu quả trong việc tạo miễn dịch thụ động chống bệnh than. Tham khảo, ví dụ, Patent Mỹ số 7,442,373; R. Sawada-Hirai và cộng sự, J Immune Based Ther Vaccines. 2004; 2: 5. (on-line 2004 May 12). Các kháng thể trung hòa độc tố kháng bệnh than khác nữa được mô tả và/hoặc có thể được tạo ra. Tương tự, các kháng thể trung hòa kháng vi khuẩn và/hoặc độc tố vi khuẩn khác có thể được sử dụng để tạo ra cấu trúc thể kháng sinh bệnh phân phối AAV như được mô tả ở đây.

Các bệnh truyền nhiễm khác có thể gây ra bởi nấm trong không khí bao gồm, ví dụ, loài *Aspergillus*, *Absidia corymbifera*, *Rhizopus stolonifer*, *Mucor plumbeus*, *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Blastomyces dermatitidis*, *Coccidioides immitis*, loài *Penicillium*, *Micropolyspora faeni*, *Thermoactinomyces vulgaris*, *Alternaria alternate*, loài *Cladosporium*, *Helminthosporium*, và loài *Stachybotrys*.

Ngoài ra, việc gây miễn dịch thụ động có thể được sử dụng để phòng ngừa lây nhiễm nấm (ví dụ, nấm bàn chân), bệnh ecpet mắng tròn, hoặc virus, vi khuẩn, ký sinh trùng, nấm, và các mầm bệnh khác mà có thể được truyền bằng cách tiếp xúc trực tiếp. Ngoài ra, nhiều tình trạng bệnh mà ảnh hưởng đến vật nuôi trong nhà, gia súc và thú nuôi khác, và các động vật khác. Ví dụ, ở chó, việc lây nhiễm đường hô hấp trên do sinonasaspergillosis ở chó gây ra bệnh đáng kể. Ở mèo, các bệnh đường hô hấp trên hoặc phức bệnh hô hấp ở mèo xuất phát trong mũi gây ra tăng nặng bệnh và chết nếu để không điều trị. Gia súc dễ nhiễm bệnh viêm thanh khí quản truyền nhiễm ở bò (thường gọi là IBR hoặc mũi đỏ) là bệnh do virus truyền nhiễm, cấp của gia súc. Ngoài ra, thú nuôi nhiễm virus hợp bào hô hấp ở bò (BRISV) mà gây ra các bệnh hô hấp từ nhẹ đến nặng và có thể làm suy yếu sức đề kháng với các bệnh khác. Các mầm bệnh và bệnh khác nữa sẽ là hiện

nhiên đối với một trong số người có hiểu biết trung bình trong lĩnh vực. Tham khảo, ví dụ, US 5,811,524, mà mô tả việc tạo ra kháng thể trung hòa kháng virus hợp bào hô hấp (RSV). Các kỹ thuật được mô tả ở đây có thể áp dụng cho các mầm bệnh khác. Kháng thể này có thể được sử dụng nguyên vẹn hoặc các trình tự của chúng (câu) được biến đổi để tạo ra cấu trúc kháng thể trung hòa nhân tạo hoặc tái tổ hợp. Phương pháp này được mô tả [tham khảo, ví dụ, WO 2010/13036; WO 2009/115972; WO 2010/140114].

Globulin miễn dịch kháng khối u như được mô tả ở đây có thể nhắm đến yếu tố tăng trưởng thượng bì ở người (HER), như HER2. Ví dụ, trastuzumab là IgG1 kappa tái tổ hợp, kháng thể người hóa đơn dòng mà liên kết chọn lọc với ái lực cao trong thử nghiệm trên cơ sở tế bào ($K_d = 5 \text{ nM}$) với vùng ngoại bào của protein thụ thể yếu tố tăng trưởng biểu bì ở người. Sản phẩm được bán trên thị trường trong nuôi cấy tế bào CHO. Tham khảo, ví dụ, <http://www.drugbank.ca/drugs/DB00072>. Các trình tự axit amin của chuỗi nhẹ 1 và 2 và chuỗi nặng 1 và 2 của trastuzumab, cũng như các trình tự thu được từ nghiên cứu về cấu trúc tia x của trastuzumab, được cấp trên cơ sở dữ liệu này với số truy cập DB00072, mà các trình tự được kết hợp vào đây bằng cách viện dẫn. Tham khảo, tương tự, 212-Pb-TCMC-trastuzumab [Areva Med, Bethesda, MD]. Kháng thể mong muốn khác bao gồm, ví dụ, pertuzumab, kháng thể người hóa đơn dòng tái tổ hợp mà nhắm đích vùng dime hóa ngoại bào (Vùng phụ II) của protein thụ thể yếu tố tăng trưởng biểu bì ở người 2 (HER2). Nó bao gồm hai chuỗi nặng và hai chuỗi nhẹ mà có 448 và 214 gốc tương ứng. FDA đã phê chuẩn ngày 8 tháng sáu năm 2012. Các trình tự axit amin của chuỗi nặng và chuỗi nhẹ của nó được đề xuất, ví dụ, trong www.drugbank.ca/drugs/DB06366 (các từ đồng nghĩa bao gồm 2C4, MOAB 2C4, kháng thể đơn dòng 2C4, và rhuMAb-2C4) trên cơ sở dữ liệu này ở số truy cập DB06366. Ngoài HER2, các đích HER khác có thể được chọn.

Ví dụ, MM-121/SAR256212 là kháng thể đơn dòng của người đầy đủ mà nhắm đích thụ thể HER3 [Merrimack's Network Biology] và mà đã được báo cáo là hữu dụng trong điều trị ung thư phổi không tiểu bào (NSCLC), ung thư vú và ung thư buồng trứng. SAR256212 là kháng thể đơn dòng của người đầy đủ điều tra mà nhắm đích thụ thể HER3 (ErbB3) [Sanofi Oncology]. Kháng thể kháng Her3/EGFR khác là RG7597 [Genentech], được mô tả là hữu dụng với ung thư đầu và cổ. Kháng thể khác, margetuximab (or MGAH22), kháng thể đơn dòng tối ưu hóa Fc, thể hệ tiếp theo (mAb) mà nhắm đích HER [MacroGenics], có thể cũng được sử dụng.

Theo cách khác, các chất đánh dấu bề mặt tế bào biểu mô của người khác và/hoặc các thụ thể u hoặc kháng thể khác có thể được nhắm đích. Các ví dụ về đích chất đánh dấu bề mặt tế bào khác bao gồm, ví dụ, 5T4, CA-125, CEA (ví dụ, được nhắm đích bởi labetuzumab), CD3, CD19, CD20 (ví dụ, được nhắm đích bởi rituximab), CD22 (ví dụ, được nhắm đích bởi epratuzumab hoặc veltuzumab), CD30, CD33, CD40, CD44, CD51 (còn gọi là integrin $\alpha_v\beta_3$), CD133 (ví dụ, tế bào u nguyên bào thần kinh đệm), CTLA-4 (ví dụ, Ipilimumab được dùng để điều trị, ví dụ, u nguyên bào thần kinh)), Chemokin (Mô típ C-X-C) thụ thể 2 (CXCR2) (được biểu hiện trong các vùng khác nhau trong não; ví dụ, kháng thể kháng CXCR2 (ngoại bào) #ACR-012 (Alomene Labs)); EpCAM, protein hoạt hóa nguyên bào sợi (FAP) [tham khảo, ví dụ, WO 2012020006 A2, ung thư não], folat thụ thể alpha (ví dụ, u não tủy ở trẻ em, ung thư đầu và cổ), thụ thể yếu tố tăng trưởng nguyên bào sợi 1 (FGFR1) (tham khảo, ví dụ, WO2012125124A1 để bàn về việc điều trị ung thư bằng kháng thể kháng FGFR1), FGFR2 (tham khảo, ví dụ, các kháng thể được mô tả trong WO2013076186A và WO2011143318A2), FGFR3 (tham khảo, ví dụ, các kháng thể được mô tả trong US8187601 và WO2010111367A1), FGFR4 (tham khảo, ví dụ, các kháng thể kháng FGFR4 được mô tả trong WO2012138975A1), yếu tố tăng trưởng tế bào gan (HGF) (tham khảo, ví dụ, các kháng thể trong WO2010119991A3), integrin $\alpha_5\beta_1$, thụ thể IGF-1, gangliosit GD2 (tham khảo, ví dụ, các kháng thể được mô tả trong

WO2011160119A2), gangiolosit GD3, glycoprotein NMB xuyên màng (GPNMB) (kết hợp với u thần kinh đệm, trong số những cái khác và đích của kháng thể glembatumumab (CR011), dịch nhầy, MUC1, phosphatidylserin (ví dụ, được nhắm đích bởi bavituximab, Peregrine Pharmaceuticals, Inc], tế bào ung thư biểu mô tuyến tiền liệt, PD-L1 (ví dụ, nivolumab (BMS-936558, MDX-1106, ONO-4538), gG4 người đầy đủ, ví dụ, ung thư tế bào hắc tố di căn], thụ thể yếu tố tăng trưởng có nguồn gốc tiểu cầu, alpha (PDGFR α) hoặc CD140, glycoprotein 72 (TAG-72) liên quan đến khối u, tenascin C, thụ thể yếu tố tiêu diệt bướu (TNF) (TRAIL-R2), yếu tố tăng trưởng nội mô mạch (VEGF)-A (ví dụ, được nhắm đích bởi bevacizumab) và VEGFR2 (ví dụ, được nhắm đích bởi ramucirumab).

Các kháng thể khác và đích của chúng bao gồm, ví dụ, APN301 (hu14.19-IL2), kháng thể đơn dòng [ung thư tế bào hắc tố và u nguyên bào thần kinh ở trẻ em, ApeironBiologics, Vienna, Austria]. Tham khảo, tương tự, ví dụ, kháng thể đơn dòng, 8H9, mà được mô tả là hữu dụng để điều trị các khối u rắn, bao gồm ung thư não di căn. Kháng thể đơn dòng 8H9 là kháng thể IgG1 chuột có tính đặc hiệu đối với kháng nguyên B7H3 [United Therapeutics Corporation]. Kháng thể chuột này có thể được người hóa. Các cấu trúc globulin miễn dịch khác nữa nhắm đích kháng nguyên B7-H3 và/hoặc kháng nguyên B7-H4 có thể được sử dụng theo sáng chế. Kháng thể khác là S58 (u nguyên bào thần kinh, kháng-GD2). Cotara™ [Peregrine Pharmaceuticals] là kháng thể đơn dòng được mô tả để điều trị u nguyên bào đệm tái hồi. Các kháng thể khác có thể bao gồm, ví dụ, avastin, ficlatuzumab, medi-575, và olaratumab. Các cấu trúc globulin miễn dịch khác nữa hoặc các kháng thể đơn dòng có thể được chọn để dùng trong sáng chế. Tham khảo, ví dụ, Medicines in Development Biologics, Báo cáo năm 2013, trang 1 - 87, công bố của PhRMA's Communications & Public Affairs Department. (202) 835-3460, được kết hợp vào đây bằng cách viện dẫn.

Ví dụ, chất sinh miễn dịch có thể được chọn từ nhiều họ virus khác nhau. Ví dụ các họ virus mà đáp ứng miễn dịch có thể mong muốn chống lại bao gồm,

họ picornavirut, mà bao gồm chủng rhinovirut, mà chịu trách nhiệm khoảng 50% trường hợp cảm lạnh; chủng enterovirut, mà bao gồm poliovirut, coxsackievirut, echovirut, và enterovirut người như virut bệnh viêm gan; và chủng aphthovirut, mà chịu trách nhiệm đối với bệnh tay chân miệng, cơ bản ở động vật không phải người. Trong họ picornavirut của virut, các kháng nguyên đích bao gồm VP1, VP2, VP3, VP4, và VPG. Trong họ virut khác bao gồm họ calcivirut, mà bao gồm nhóm Norwalk của virut, mà là tác nhân chủ yếu quan trọng của viêm dạ dày ruột dịch. Họ virut khác nữa mong muốn để dùng để hướng đích kháng nguyên để tạo ra đáp ứng miễn dịch ở người và động vật không phải người là họ togavirut, mà bao gồm chủng alphavirut, mà bao gồm virut Sindbis, virut RossRiver, và viêm não ngựa Venezuela, phương Đông & phương Tây, và rubivirut, bao gồm Rubellvirut. Họ flaviviridae bao gồm virut sốt xuất huyết, sốt vàng, viêm não Nhật bản, viêm não St. Louis và viêm não do ve truyền. Các kháng nguyên đích khác có thể được tạo ra từ viêm gan C hoặc họ coronavirut, mà bao gồm nhiều virut không phải ở người như virut viêm phế quản truyền nhiễm (gà), virut viêm dạ dày ruột truyền nhiễm trên heo (lợn), virut viêm não tủy ngưng kết hồng cầu ở lợn (lợn), virut viêm phúc mạc truyền nhiễm ở mèo (mèo), coronavirut gây bệnh đường ruột ở mèo (mèo), coronavirut ở chó (chó), và coronavirut hô hấp ở người, mà có thể gây ra cảm lạnh và/hoặc viêm gan không A, B hoặc C. Trong họ coronavirut, các kháng nguyên đích bao gồm E1 (còn gọi là M hoặc protein nền), E2 (còn gọi là S hoặc protein Spike), E3 (còn gọi là HE hoặc hemagglutinin-esteraza) glycoprotein (không có mặt trong tất cả coronavirut), hoặc N (vỏ protein bọc nhân). Các kháng nguyên khác nữa có thể được nhắm đích chống họ rhabdovirut, mà bao gồm chủng vesiculovirut (ví dụ, virut gây viêm miệng có mụn nước), và chủng lyssavirut (ví dụ, bệnh dại).

Trong họ rhabdovirut, kháng nguyên thích hợp có thể thu được từ protein G hoặc protein N. Họ filoviridae, mà bao gồm virut sốt xuất huyết như Marburg và Ebolvirut, có thể là nguồn kháng nguyên thích hợp. Họ paramyxovirut bao gồm virut parainfluenza tuýp 1, virut parainfluenza tuýp 3, virut parainfluenza tuýp 3

ở bò, rubulavirut (virut quai bị), virut parainfluenza tuýp 2, virut parainfluenzatuýp 4, virut bệnh Newcastle (gà), bệnh dịch hạch gia súc, morbillivirut, mà bao gồm bệnh sởi và bệnh sài sốt chó, và pneumovirut, mà bao gồm virut hợp bào hô hấp. Virut cúm được phân loại họ orthomyxovirut và là nguồn kháng nguyên thích hợp (ví dụ, protein HA, protein N1). Họ bunyavirut bao gồm chủng bunyavirut (viêm não California, La Crosse), phlebovirut (sốt thung lũng Rift), hantavirut (puremala là virut sốt hemahagin), nairovirut (bệnh cừu Nairobi) và nhiều bungavirut không xác định khác. Họ arenavirut cung cấp nguồn kháng nguyên chống LCM và virut sốt Lassa. Họ reovirut bao gồm chủng reovirut, rotavirut (mà gây ra viêm dạ dày ruột cấp ở trẻ em), orbivirut, và cultivirut (sốt ve Colorado, Lebombo (người), thoái hóa não ngựa, bệnh lưỡi xanh).

Họ retrovirut bao gồm phân họ oncorivirinal mà bao gồm các bệnh thú y và người như suy giảm bạch cầu ở mèo, HTLV I và HTLV II, lentivirinal (mà bao gồm virut gây suy giảm miễn dịch ở người (HIV), virut gây suy giảm miễn dịch ở khỉ (SIV), virut gây suy giảm miễn dịch ở mèo (FIV), virut gây thiếu máu lây nhiễm ở ngựa, và spumavirinal). Trong số lentivirut, nhiều kháng nguyên thích hợp được mô tả và có thể được chọn sẵn làm đích. Các ví dụ của kháng nguyên HIV và SIV thích hợp bao gồm, không giới hạn các protein gag, pol, Vif, Vpx, VPR, Env, Tat, Nef, và Rev, cũng như nhiều đoạn khác nhau của chúng. Ví dụ, đoạn thích hợp của protein Env có thể bao gồm bất kỳ trong các khối phụ như gp120, gp160, gp41, hoặc đoạn nhỏ hơn của chúng, ví dụ, có chiều dài bằng ít nhất khoảng 8 axit amin. Tương tự, đoạn của protein tat có thể được chọn. [Tham khảo, Patent Mỹ số 5,891,994 và Patent Mỹ số 6,193,981.] Tham khảo, tương tự, các protein HIV và SIV được mô tả trong D.H. Barouch và cộng sự, *J. Virol.*, 75(5):2462-2467 (March 2001), và R.R. Amara, và cộng sự, *Science*, 292:69-74 (6 April 2001). . Trong ví dụ khác, protein gây miễn dịch HIV và/hoặc SIV hoặc peptit có thể được sử dụng để tạo ra protein dung hợp hoặc các phân tử gây miễn dịch khác. Tham khảo, ví dụ, protein dung hợp HIV-1 Tat và/hoặc Nef và các cơ

chế gây miễn dịch được mô tả trong WO 01/54719, công bố ngày 2 tháng tám năm 2001, và WO 99/16884, công bố ngày 8 tháng tư năm 1999. Sáng chế không bị giới hạn ở protein gây miễn dịch HIV và/hoặc SIV hoặc peptit được mô tả ở đây. Ngoài ra, nhiều cải biến của các protein này được mô tả hoặc có thể được tạo sẵn bởi một trong số các người có hiểu biết trung bình trong lĩnh vực. Tham khảo, ví dụ, protein gag biến đổi mà được mô tả trong Patent Mỹ số 5,972,596.

Họ papovavirut bao gồm phân họ polyomavirut (các virus BKU và JCU) và phân họ papillomavirut (liên quan đến các bệnh ung thư hoặc sự tiến triển ác tính u nhú). Họ adenovirut bao gồm các virus (EX, AD7, ARD, O.B.) mà gây ra các bệnh hô hấp và/hoặc viêm ruột. Họ parvovirut bao gồm parvovirut ở mèo (bệnh viêm ruột ở mèo), panleucopeniavirut ở mèo, parvovirut ở chó, và parvovirut ở lợn. Họ herpesvirut bao gồm phân họ alphaherpesvirinae, mà bao gồm loại simplexvirut (HSVI, HSVII), varicellovirut (pseudorabies, varicella zoster) và phân họ betaherpesvirinae, mà bao gồm loại cytomegalovirut (HCMV, muromegalovirut) và phân họ gammaherpesvirinae, mà bao gồm loài lymphocryptovirut, EBV (u lympho Burkitts), virus viêm mũi truyền nhiễm, virus gây bệnh Marek, và rhadinovirut. Họ poxvirut bao gồm phân họ chordopoxvirinae, mà bao gồm giống orthopoxvirut (Variola (Bệnh đậu mùa) và Vaccinia (Cowpox)), parapoxvirut, avipoxvirut, capripoxvirut, leporipoxvirut, suipoxvirut, và phân họ entomopoxvirinae. Họ hepadnavirut bao gồm virus viêm gan B. Một virus không được phân loại mà có thể là nguồn thích hợp của kháng nguyên là virus viêm gan delta. Các nguồn virus khác có thể bao gồm virus gây bệnh Gumboro ở gia cầm và virus gây hội chứng rối loạn hô hấp và sinh sản ở lợn. Họ alphavirut bao gồm virus gây viêm động mạch ngựa và nhiều virus gây viêm não khác nhau.

Các đích gây mầm bệnh khác cho các kháng thể có thể bao gồm, ví dụ, vi khuẩn, nấm, vi sinh vật ký sinh hoặc ký sinh trùng đa bào mà lây nhiễm cho người và động vật có xương sống không phải người, hoặc từ tế bào ung thư hoặc tế bào

u. Các ví dụ của mầm bệnh vi khuẩn bao gồm cầu khuẩn gram dương sinh bệnh bao gồm phế cầu khuẩn; tụ cầu khuẩn; và liên cầu khuẩn. Cầu khuẩn gram âm sinh bệnh bao gồm cầu khuẩn màng não; lậu cầu khuẩn. Vi khuẩn gram âm ruột sinh bệnh bao gồm enterobacteriaceae; pseudomonas, acinetobacteria và eikenella; bệnh nhiễm khuẩn Malleomyces pseudomallei; salmonella; shigella; haemophilus; moraxella; *H. ducreyi* (mà gây bệnh hạ cam); brucella; *Fransisella tularensis* (mà gây bệnh tularemia); yersinia (pasteurella); streptobacillusmoniliformis và spirillum; vi khuẩn Gram dương bao gồm listeria monocytogenes; erysipelothruxrhusiopathiae; *Corynebacterium diphtheria* (bệnh bạch cầu); bệnh tả; *B. anthracis* (bệnh than); donovanosis (u hạt bẹn); và bartonellosis. Các bệnh gây ra bởi vi khuẩn kỵ khí sinh bệnh bao gồm bệnh uốn ván; chứng ngộ độc thịt; clostridia; bệnh lao; leprosy; và mycobacteria khác. Các bệnh do xoắn khuẩn sinh bệnh bao gồm bệnh giang mai; bệnh Treponema: bệnh ghẻ cóc, sốt pinta và dịch giang mai địa phương; và bệnh leptospirosis. Các bệnh lây nhiễm do vi khuẩn của mầm bệnh cao hơn và nấm gây bệnh bao gồm bệnh do khuẩn tia; bệnh Nocardia; bệnh cryptococcosis, bệnh nấm chồi, bệnh Histoplasme và bệnh nấm coccidioides; bệnh nấm candidia, bệnh do nấm aspergillus, và bệnh nấm Mucor; bệnh nấm sporotrichum; bệnh nấm paracoccidioidomycosis, bệnh nấm petriellidiosis, bệnh nấm torulopsosis, bệnh nấm Mycetoma và bệnh nấm sâu chromomycosis; và bệnh da do nấm sợi (dermatophytosis). Các lây nhiễm do Rickettsia bao gồm sốt phát ban, sốt màng não miền núi, bệnh sốt Q, và bệnh phát ban do Rickettsia. Các ví dụ về lây nhiễm Mycoplasma và Chlamydia bao gồm: bệnh viêm phổi do mycoplasma pneumoniae; bệnh hội xoài; bệnh virus vệt; và nhiễm Chlamydia niệu đạo. Sinh vật nhân chuẩn gây bệnh bao gồm động vật nguyên sinh sinh bệnh và giun sán và lây nhiễm tạo ra từ đó bao gồm: bệnh amip; bệnh sốt rét; bệnh leishmania; bệnh do trypanosoma; bệnh do toxoplasma; *Pneumocystis carinii*; *Trichans*; *Toxoplasma gondii*; chứng nhiễm ký sinh trùng babesia; bệnh do giardia; nhiễm giun xoắn; giun chỉ; sán máng; giun tròn; sán lá hoặc sán gan; và sán dây (sán xơ mít).

Nhiều trong số sinh vật và/hoặc độc tố này tạo ra từ đó đã được xác định bởi Trung tâm kiểm soát bệnh [(Centers for Disease Control - CDC), Bộ y tế và dịch vụ con người, Mỹ], là các chất mà có tiềm năng để dùng trong tấn công sinh học. Ví dụ, một số trong các tác nhân sinh học, bao gồm, *Bacillus anthracis* (bệnh than), *Clostridium botulinum* và độc tố của nó (chứng ngộ độc thịt), *Yersinia pestis* (bệnh dịch hạch), variola major (bệnh đậu mùa), *Francisellatularensis* (bệnh tularemia), và sốt xuất huyết [filovirut (ví dụ, Ebola, Marburg], và arenavirut [ví dụ, Lassa, Machupo]), tất cả trong đó hiện được xếp vào các tác nhân Nhóm A; *Coxiella burnetti* (Bệnh sốt Q); các loài *Brucella* (bệnh sảy thai truyền nhiễm), *Burkholderia mallei* (bệnh loét mũi truyền nhiễm), *Burkholderia pseudomallei* (bệnh Whitmore), *Ricinus communis* và độc tố của nó (độc tố rixin), *Clostridium perfringens* và độc tố của nó (độc tố epsilon), các loài *Staphylococcus* và các độc tố của chúng (độc tố trong ruột B), *Chlamydia psittaci* (bệnh virut vẹt), nguy cơ mất nước (ví dụ, *Vibrio bệnh tả*, *Cryptosporidium parvum*), sốt phát ban (*Richettsia powazekii*), và viêm não do virut (alphavirut, ví dụ, viêm não ngựa Venezuela; viêm não ngựa phương Đông; viêm não ngựa phương Tây); tất cả chúng hiện được xếp vào các tác nhân Nhóm B; và virut và hantavirut Nipan, hiện được xếp vào các tác nhân Nhóm C. Ngoài ra, các sinh vật khác, mà được xếp nhóm giống hoặc khác, có thể được xác định và/hoặc sử dụng cho mục đích này trong tương lai. Cần hiểu rằng các vector virut và các cấu trúc khác được mô tả ở đây là hữu dụng cho các kháng nguyên đích từ các sinh vật, virut này, các độc tố của chúng hoặc các sản phẩm khác từ chúng, mà sẽ phòng ngừa và/hoặc điều trị nhiễm hoặc phản ứng bất lợi khác với các tác nhân sinh học.

Ví dụ thực hiện sáng chế

Các ví dụ sau đây chỉ nhằm mục đích minh họa và không làm giới hạn sáng chế được mô tả ở đây.

Ví dụ 1: Tạo ra các vector chứa cat-xet đồng biểu hiện của kháng thể chiều dài đầy đủ

Hàng loạt các cis-plasmit được điều chế để dùng trong việc tạo ra phân tử virut AAV chứa phân tử axit nucleic để phân phối đến tế bào đích chủ. Các phân tử axit nucleic chứa các trình tự ITR ở đầu 5' và 3' của AAV2 ở mỗi đầu cuối, vùng tăng cường CMV chung bên cạnh hai cat-xet biểu hiện theo hai hướng đối lập, trong đó cat-xet biểu hiện thứ nhất được điều khiển bởi vùng khởi đầu CMV tối thiểu thứ nhất và cat-xet biểu hiện thứ hai được kiểm soát bởi vùng khởi đầu CMV tối thiểu thứ hai. Tất cả các trình tự nằm giữa các ITR của AAV2 được tổng hợp mới bởi các nhà cung cấp trên thị trường (GeneArt). Tất cả các trình tự mã hóa cho vùng biến đổi globulin miễn dịch ở cạnh các enzym giới hạn duy nhất để tạo ra việc xáo trộn thích hợp của vùng biến đổi mong muốn. Để tạo ra các cấu trúc với trình tự chuỗi nhẹ khác loại (kgl), trình tự mã hóa mã hóa chuỗi nhẹ dòng gốc (IGKV4-1*01) mà được tổng hợp mới và được sử dụng để thay thế Trình tự chuỗi nhẹ biến đổi FI6.

Việc xáo trộn đồng biểu hiện của kháng thể ví dụ được thể hiện trên Hình 2. Việc xáo trộn này gồm ở bên trái của vùng tăng cường cat-xet biểu hiện thứ nhất mà gồm, từ phải sang trái, vùng khởi đầu tối thiểu CMV, trình tự đoạn dẫn IL2 khác loại liên kết với vùng biến đổi chuỗi nặng (VH) của kháng thể kháng TSG101 (1A6), vùng CH'1, và vùng CH'2-3 mà được tối ưu hóa để biểu hiện ở người, và polyA tổng hợp. Nằm ở bên phải của vùng tăng cường là vùng khởi đầu tối thiểu CMV, trình tự đoạn dẫn IL2 khác loại, vùng biến đổi chuỗi nhẹ và vùng hằng định chuỗi nhẹ của FI6k2 (kháng thể kháng cúm), vị trí phân cắt furin, liên kết 2a từ virut bệnh chân tay miệng, trình tự đoạn dẫn IL2, FI6v3 VH, CH1, CH2-3, và trình tự polyA ngăn thymidin kinaza. Các chỉ định CH đề cập đến alen G1m17,1 của kháng thể đã biết.

SEQ ID NO: 1 đề xuất các trình tự của vùng hằng định FI6. Các trình tự axit amin của chuỗi nhẹ axit amin FI6 được nêu ở SEQ ID NO: 2.

Cis-plasmit của Hình 2 được sử dụng trong phương pháp chuyển nhiễm ba lần như mô tả ở trước trong, ví dụ, công bố đơn Patent Mỹ số 12/226,558, để tạo

ra các vector AAV8 và AAV9 mà được sử dụng trong các nghiên cứu tiếp theo được mô tả ở đây. Plasmid thu được, pN509_ACE Fi6-1A6 MAB_p3160, dài 7722 bp, trình tự trong đó được nêu ở SEQ ID NO: 3, được kết hợp ở đây bằng cách viện dẫn cùng với các dấu hiệu của nó. Các trình tự được mã hóa cho chuỗi nhẹ biến đổi FI6 (VL) [SEQ ID NO:4], chuỗi nặng biến đổi FI6 [SEQ ID NO: 5], CH1 (SEQ ID NO: 6), CH2-3 [SEQ ID NO: 7] cũng được đề xuất.

Cis-plasmid đồng biểu hiện của kháng thể tương tự được tạo ra bằng cách tách dòng lại kháng thể cúm mùa (CR8033) hoặc kháng thể cúm dịch (C05), hoặc kháng thể kháng M2e (TCN-032 thay cho vùng biến đổi chuỗi nặng 1A6 trên Hình 2 sử dụng các điểm giới hạn duy nhất định vị trước mà cho phép dễ dàng xáo trộn các vùng biến đổi. Các cis-plasmid này lần lượt được dùng trong chuyển nhiễm ba lần (ví dụ, được thực hiện như được mô tả trong đơn Patent Mỹ số 12/226,588) để tạo ra các vector AAV8 và AAV9 được dùng cho các nghiên cứu tiếp theo. Các trình tự đối với xáo trộn pN510_ACE Fi6-C05 MAB được nêu ở SEQ ID NO:8; trình tự axit amin của chuỗi nhẹ biến đổi được nêu ở SEQ ID NO: 9, chuỗi nhẹ hằng định được nêu ở SEQ ID NO: 10, chuỗi nặng biến đổi FI6 được nêu ở SEQ ID NO: 11, CH1 được nêu ở SEQ ID NO:12 và CH2-3 được nêu ở SEQ ID NO: 13.

Các trình tự đối với xáo trộn pN514_ACE Fi6-C05 MAB được nêu ở SEQ ID NO:19; trình tự axit amin của chuỗi nhẹ hằng định được nêu ở SEQ ID NO: 20, chuỗi nặng biến đổi FI6 được nêu ở SEQ ID NO: 21, CH1 được nêu ở SEQ ID NO:22 và CH2-3 được nêu ở SEQ ID NO: 23. Các vùng qua lại này đến lượt mình được sử dụng để tạo ra các vector AAV8 và AAV9 mà được sử dụng cho các nghiên cứu tiếp theo.

Ví dụ 2: Xác định các sản phẩm được biểu hiện từ các vector AAV8 đồng biểu hiện kháng thể đơn dòng (MAB) F16 VÀ MAB IA6

Hàng loạt các thử nghiệm ELISA được thực hiện để xác định mức độ biểu hiện và đánh giá sự liên kết của FI6 MAB được đồng biểu hiện với IA6 MAB từ

cis plasmit được tạo ra như được mô tả trong ví dụ 1 sau khi chuyển nhiễm vào tế bào HEK 293. Peptit TSG101 được tổng hợp sử dụng hóa học f-Moc bởi Mimotopes. Tất cả kháng nguyên cúm thu được từ nhà cung cấp trên thị trường, ImmuneTechnologies, Inc. ProteinA được mua từ Sigma-Aldrich và được sử dụng để theo dõi biểu hiện của tổng IgG1 ở người. Việc phát hiện IgG1 người trong dịch nổi nuôi cấy mô được đo bằng hoặc là ELISA đặc hiệu kháng nguyên hoặc bắt proteinA. Tấm ELISA liên kết cao được phủ với 2 $\mu\text{g/ml}$ protein HA hoặc peptit, hoặc với 5 $\mu\text{g/ml}$ proteinA được pha loãng trong PBS và được ủ qua đêm ở 4°C. Các giếng được rửa 5-8 lần và được chặn bằng 1 mM EDTA, 5% PBS tro nhiệt, 0,07% Tween 20 trong PBS trong một giờ ở nhiệt độ phòng. Dịch nổi nuôi cấy mô được thêm vào các tấm ở các độ pha loãng khác nhau trong các mẫu lặp và được ủ ở 37°C trong một giờ. Các tấm được rửa, được chặn, và kháng thể IgG dê kháng người Affinipures liên hợp Bio-SP (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA, USA) được thêm ở tỷ lệ pha loãng 1:10.000. Sau một giờ, các tấm được rửa và peroxidaza cải ngựa liên hợp streptavidin (HRP) được thêm ở tỷ lệ pha loãng 1:30.000. Sau một giờ, các tấm được rửa 3,3',5,5'-tetramethylbenzidin (TMB) được thêm. Phản ứng được dừng sau 30 phút ở nhiệt độ phòng sử dụng axit sulfuric 2N và các tấm được đọc ở 450 nm sử dụng thiết bị đọc tấm BioTek μQuant (Winooski, VT, USA).

Như được mong đợi, không có liên kết được thấy của FI6 với peptit TSG101, HA (B/Malaysia/2506/2/004), hoặc HA (chỉ vùng đầu của chủng cúm A/Brisbane/59/2007). Việc liên kết FI6 được quan sát thấy cho cùng chủng cúm khi HA chiều dài đầy đủ có mặt, cũng như cho chủng cúm HA(dTM)(A/Beijing/01/2009, H1N1)). Như được mong đợi, việc liên kết FI6 cũng quan sát thấy đối với Protein A.

Theo báo cáo được công bố, FI6 được tạo ra theo phương pháp trong lĩnh vực liên kết với liên kết với HA chiều dài đầy đủ và thân HA, nhưng không liên

kết với vùng chỉ đầu. Các dữ liệu này chứng minh rằng kháng thể đơn dòng FI6 được đồng biểu hiện duy trì đặc điểm liên kết đặc trưng của nó.

Ví dụ 3: Xác định các sản phẩm được biểu hiện từ các vectơ AAV8 đồng biểu hiện kháng thể đơn dòng (MAB) F16 và cúm dịch cúm MAB C05

Khả năng phát hiện khác nhau của hai kháng thể đơn dòng khác nhau được đánh giá trong thử nghiệm bắt. Các kháng thể đơn dòng FI6 và C05 được đồng biểu hiện từ cis-plasmit được điều chế như được mô tả trong ví dụ 1 và được chuyển nhiễm vào tế bào HEK293 được đánh giá cho việc liên kết. FI6 được mong đợi là liên kết với HA chiều dài đầy đủ và thân HA, nhưng không liên kết với vùng chỉ đầu. Các kết quả của nghiên cứu liên kết được minh họa trên Hình 3 chứng minh rằng các kháng thể được đồng biểu hiện duy trì liên kết đặc trưng của chúng. Cụ thể hơn là, đặc trưng liên kết với HA chiều dài đầy đủ và thân HA đặc thù của FI6 được quan sát và đặc trưng liên kết với HA và HA chỉ đầu (không thân) của C05 cũng được quan sát. Các thử nghiệm ELISA được thực hiện như được mô tả trong ví dụ 2.

Ví dụ 4: Xác định các sản phẩm được biểu hiện từ các vectơ AAV8 đồng biểu hiện kháng thể đơn dòng (MAB) F16 và MAB chiều dài đầy đủ thứ hai

Chuột RAG KO đực 6-8 tuần tuổi (The Jackson Laboratory Bar Harbor, ME, USA) được ở trong các điều kiện không có mầm bệnh tại Phòng thí nghiệm nghiên cứu dịch mã của Đại học Pennsylvania. Toàn bộ quy trình và thủ tục cho động vật được phê chuẩn bởi ủy ban chăm sóc và sử dụng động vật nghiên cứu (the Institutional Animal Care and Use Committee). Chuột được làm chết bằng cách làm ngạt bằng cacbon đioxit và xác nhận chết bằng cách làm trật khớp cổ. Để dùng vectơ, chuột được gây mê bằng hỗn hợp của 70 mg/kg trọng lượng cơ thể ketamin và 7 mg/kg trọng lượng cơ thể xylazin bằng cách tiêm màng bụng (IP). Các vectơ được pha loãng trong nước muối đệm phosphat (PBS) và các lần tiêm IM được thực hiện sử dụng ống tiêm Hamilton. Huyết thanh được thu hàng tuần qua máu hốc mắt.

Việc phát hiện IgG1 người trong dịch nổi nuôi cấy mô được đo bằng thử nghiệm ELISA bắt protein A. Các tấm ELISA liên kết cao được phủ với 5 µg/ml proteinA được pha loãng trong PBS và được ủ qua đêm ở 4°C. Các giếng được rửa 5-8 lần và được chặn bằng EDTA 1 mM, PBS 5% tro nhiệt, Tween 20 0,07% trong PBS. Các mẫu huyết thanh chuột được làm tro nhiệt và được thêm vào các tấm ở các độ pha loãng khác nhau trong các mẫu lặp và được ủ ở 37°C trong một giờ. Các tấm được rửa, được chặn, và kháng thể IgG dê kháng người Affinipures liên hợp Bio-SP (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA, USA) được thêm ở tỷ lệ pha loãng 1:10.000. Sau một giờ, các tấm được rửa và được ủ với peroxidaza cải ngựa liên hợp streptavidin (HRP) ở độ pha loãng 1:30.000. Sau một giờ, các tấm được rửa 3,3',5,5'-tetrametylbenzidin (TMB) được thêm. Phản ứng được dừng sau 30 phút ở nhiệt độ phòng sử dụng axit sulfuric 2N và các tấm được đọc ở 450 nm sử dụng thiết bị đọc tấm BioTek µQuant (Winooski, VT, USA).

Hình 5 minh họa mức độ biểu hiện hệ thống cho tổng số IgG1 người ở chuột được cho dùng vectơ AAV đồng biểu hiện FI6 với kháng thể IA6. Chuột được tiêm trong cơ với các liều lượng 1×10^{11} bản sao bộ gen (GC) hoặc 1×10^{10} GC. Mức độ biểu hiện được đánh giá vào ngày thứ 7, 15, 21, 28, 34, 42, 49 và 56 và được đo ở nồng độ bằng microgram/mL. Việc tăng biểu hiện phụ thuộc liều được quan sát thấy.

Ví dụ 5: Xác định các sản phẩm được biểu hiện từ các vectơ AAV8 đồng biểu hiện kháng thể đơn dòng (MAB) F16 và ba kháng thể đơn dòng chiều dài đầy đủ khác

Các bảng dưới đây thể hiện mức độ biểu hiện ở chuột được cho dùng vectơ AAV đồng biểu hiện FI6 với kháng thể đơn dòng CR8033, C05, hoặc 1A6 chiều dài đầy đủ. Chuột được làm bất hoạt (KO) RAG được tiêm trong cơ với các liều dùng bằng 1×10^{11} bản sao bộ gen (GC) hoặc 1×10^{10} GC như được mô tả trong ví dụ trước. Mức độ biểu hiện được đánh giá hàng tuần vào các ngày thứ 7, 15,

21, 28, 34, 42, và 49 và được đo ở nồng độ bằng microgram/mL. Việc tăng biểu hiện phụ thuộc liều được quan sát thấy đối với kháng thể được biểu hiện. Kháng nguyên bắt được dùng cho thử nghiệm là Protein A ELISA như được mô tả trong ví dụ trước.

Mẫu thử	Fi6v3k2 mAb + CR8033 mAb			
Liều lượng	1,00 x 10 ¹¹		1,00 x 10 ¹⁰	
	Trung bình	stdev.	Trung bình	stdev.
Ngày 0	0,00	0,00	0,00	0,00
Ngày 7	2,92	0,48	0,04	0,07
Ngày 14	18,30	4,79	1,24	0,66
Ngày 21	33,69	7,45	2,09	0,88
Ngày 28	43,38	10,92	2,84	1,81
Ngày 35	66,45	16,61	4,47	1,86
Ngày 42	64,25	12,06	4,37	2,35
Ngày 49	51,36	11,90	3,57	1,52

Mẫu thử	Fi6v3k2 mAb + CO5 mAb			
Liều lượng	1,00 x 10 ¹¹		1,00 x 10 ¹⁰	
	Trung bình	stdev.	Trung bình	stdev.
Ngày 0	0,00	0,00	0,00	0,00
Ngày 7	1,73	0,42	0,00	0,00
Ngày 14	9,95	3,39	0,24	0,22
Ngày 21	24,74	11,66	0,81	0,24

Ngày 28	22,32	4,77	1,11	0,17
Ngày 35	31,67	7,93	1,53	0,28
Ngày 42	34,69	14,46	1,83	0,29
Ngày 49	26,14	5,85	1,46	0,49

Mẫu thử	Fi6v3k2 mAb + 1A6 mAb			
Liều lượng	1,00 x 10 ¹¹		1,00 x 10 ¹⁰	
	Trung bình	stdev.	Trung bình	stdev.
Ngày 0	0	0	0	0
Ngày 7	2,70	0,75	0	0
Ngày 14	5,01	0,06	1,58	,055
Ngày 21	30,16	13,31	1,71	0,52
Ngày 28	38,18	15,99	2,16	0,59
Ngày 35	55,18	18,52	4,09	1,53
Ngày 42	50,49	16,61	3,69	0,94
Ngày 49	46,66	15,59	3,73	1,09

Ví dụ 6: Hiệu quả kháng virus thu được nhờ các kháng thể chiều dài đầy đủ kép được biểu hiện từ vector AAV9 và/hoặc AAV8 đơn trong cơ

A. AAV9.BiD.FI6_CR8033mAb và thử thách cúm A

Chuột BALB/c được tiêm trong cơ với AAV9.BiD.FI6_CR8033mAb phân phối được phân phối trong cơ (IM) ở 1 x 10¹¹ GC. Hai tuần sau chuột được thử thách trong mũi với 5LD50 của chuột thích ứng PR8 (cúm A). Hình tròn thể hiện

cấu trúc AAV9 có vùng khởi đầu hai hướng biểu hiện các kháng thể đơn dòng FI6 và CR8033 tổng hợp có cùng chuỗi nhẹ khác loại. Hình vuông thể hiện đối chứng dương, tức là, AAV9 biểu hiện kháng thể đơn loại FI6 cũng được phân phối ở 1×10^{11} GC, và hình tam giác thể hiện các động vật tự nhiên. Hình 6B thể hiện sống sót sau thử thách. Việc dùng AAV9.BiD.FI6_CR8033mAb với liều lượng 1011 GC/chuột cho phép bảo vệ một phần với việc làm chậm đáng kể việc giảm khối lượng.

B. AAV9.BiD.FI6_CR8033mAb và thử thách cúm B

Đối với tiêm vector AAV9: chuột cái BALB/c được gây mê bằng cách tiêm trong cơ 100 mg/kg ketamin/ 10 mg/kg hỗn hợp xylazin trong PBS, và vector AAV9.BiD.FI6_CR8033mAb được tiêm rong cơ (IM) ở 1×10^{11} GC mỗi chuột. Vector BiD được so sánh với AAV9 biểu hiện tuýp kháng thể đơn CR8033 cũng được phân phối ở 1×10^{11} GC, và đối chứng âm (các động vật tự nhiên). Hình 7B thể hiện sống sót sau thử thách. Đối với thử thách cúm, hai tuần sau khi điều trị vector, chuột BALB/c được điều trị AAV và tự nhiên được cân và mã hóa màu đuôi, được gây mê như được mô tả ở trên, được treo bởi răng cửa giữa của chúng với các chân sau của chúng đỡ trên sàn, và đưa vào trong mũi 5LD50 of B/Lee/40 (cúm B) trong tổng khối lượng bằng 50 μ l PBS như được mô tả ở trên. Sau đó chuột được cân hàng ngày và được theo dõi các dấu hiệu của bệnh hoặc kiệt sức. Các động vật mà biểu hiện dấu hiệu đáng kể kiệt sức hoặc giảm 30% tổng trọng lượng ban đầu của chúng bị làm chết bằng cách làm ngạt bằng khí CO₂.

Hình 7A là đồ thị đường thể hiện phần trăm thay đổi về trọng lượng. Các dữ liệu này thể hiện rằng hiệu quả bảo vệ đầy đủ được tạo ra nhờ các kháng thể được biểu hiện kép ở liều lượng này. Hình 7B thể hiện sống sót sau thử thách.

C. Các vector AAV8.FI6-TCN032, AAV8.FI6-1A6, và AAV8.FI6-CR8033 được dùng IM và chuột được thích ứng thử thách cúm A PR8.

Các vector này được tạo ra như được mô tả trong ví dụ 1. Chuột RAG KO đực 6-8 tuần tuổi (The Jackson Laboratory Bar Harbor, ME, USA) được ở trong

các điều kiện không có mầm bệnh tại Phòng thí nghiệm nghiên cứu dịch mã của Đại học Pennsylvania. Toàn bộ quy trình và thủ tục cho động vật được phê chuẩn bởi ủy ban chăm sóc và sử dụng động vật nghiên cứu (the Institutional Animal Care and Use Committee). Để dùng vector, chuột được gây mê bằng hỗn hợp của 70 mg/kg trọng lượng cơ thể ketamin và 7 mg/kg trọng lượng cơ thể xylazin bằng cách tiêm màng bụng (IP). Các vector được pha loãng trong nước muối đệm phosphat (PBS) và tiêm IM được thực hiện sử dụng ống tiêm Hamilton. Huyết thanh được thu hàng tuần qua máu hốc mắt.

Phát hiện IgG1 người trong dịch nổi nuôi cấy mô được đo bằng thử nghiệm ELISA bắt proteinA. Tấm ELISA liên kết cao được phủ với 5 µg/ml proteinA được pha loãng trong PBS và được ủ qua đêm ở 4°C. Các giếng được rửa 5-8 lần và được chặn bằng EDTA 1 mM, PBS 5% trơ nhiệt, Tween 20 0,07% trong PBS. Các mẫu huyết thanh chuột được làm trơ nhiệt và được thêm vào các tấm ở các độ pha loãng khác nhau trong các mẫu lặp và được ủ ở 37°C trong một giờ. Các tấm được rửa, được chặn, và kháng thể IgG dê kháng người Affinipures liên hợp Bio-SP (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA, USA) được thêm ở tỷ lệ pha loãng 1:10.000. Sau một giờ, các tấm được rửa và được ủ với peroxidaza cải ngựa liên hợp streptavidin (HRP) ở độ pha loãng 1:30.000. Sau một giờ, các tấm được rửa 3,3',5,5'-tetramethylbenzidin (TMB) được thêm. Phản ứng được dừng sau 30 phút ở nhiệt độ phòng sử dụng axit sulfuric 2N và các tấm được đọc ở 450 nm sử dụng thiết bị đọc tấm BioTek µQuant (Winooski, VT, USA).

Xem trên Hình 8C, trên tất cả các tấm, mức độ biểu hiện được chỉ ra vào Ngày 56 sau khi cho dùng vector. Hai ngày sau khi chảy máu hốc mắt cuối cùng vào Ngày 56, chuột được cân khối lượng chuột và mã hóa màu đuôi, được gây mê như được mô tả ở trên, được treo bởi răng cửa giữa của chúng với các chân sau của chúng đỡ trên sàn, và được đưa vào trong mũi với 5LD₅₀ của chuột thích ứng PR8 (cúm A) trong tổng khối lượng bằng 50µl PBS như được mô tả ở trên. Sau

đó chuột được cân hàng ngày và được theo dõi các dấu hiệu của bệnh hoặc kiệt sức. Các động vật mà biểu hiện dấu hiệu đáng điều kiệt sức hoặc giảm 30% tổng trọng lượng ban đầu của chúng bị làm chết bằng cách làm ngạt bằng khí CO₂ và xác nhận chết bằng cách làm trật khớp cổ. Hình 8A thể hiện rằng biểu hiện hệ thống của chỉ 25 µg/ml kháng thể kháng cúm đủ để tạo ra bảo vệ trong thử thách PR8, nhưng biểu hiện bằng 0,4 µg/ml đủ để bảo vệ.

D. AAV9. FI6_IA6 mAbs và Thử thách cúm A

Vector AAV9 biểu hiện FI6 nhân tạo và adhesin miễn dịch kháng HIV, IA6, được đánh giá đối với bảo vệ chống thử thách với cúm A như được mô tả ở trên. Hình 8B thể hiện rằng biểu hiện 36,5 µg/ml của kháng thể kháng cúm đủ để tạo ra bảo vệ hoàn toàn chống lại thử thách với PR8. Hình 8C thể hiện biểu hiện 6,9 µg/ml của các kháng thể kháng cúm không đủ để bảo vệ chống lại thử thách PR8.

Ví dụ 8 – Tạo ra vector chứa hai cat-xet đồng biểu hiện adhesin miễn dịch

Sử dụng vector con thoi tương tự như được minh họa trên Hình 2, các vector chứa hai adhesin miễn dịch được tạo ra.

Theo một phương án, vector chứa các adhesin miễn dịch FI6 và C05 được tạo. Các trình tự từ plasmit mang cat-xet biểu hiện adhesin miễn dịch FI6 và CO5 được nêu ở SEQ ID NO: 36; với các trình tự mã hóa được dịch mã được nêu ở SEQ ID NO: 37 (chuỗi nặng biến đổi FI6), SEQ ID NO: 38 (chuỗi nhẹ biến đổi FI6), và SEQ ID NO: 39 (CH2-3). Các trình tự này và các đặc điểm của chúng được kết hợp bằng cách viện dẫn.

Theo phương án khác, vector chứa các adhesin miễn dịch FI6 và CR8033 được tạo. Các trình tự từ plasmit chứa các adhesin miễn dịch FI6 và CR8033 được nêu ở SEQ ID NO:40; với các trình tự mã hóa được dịch mã được nêu ở SEQ ID NO: 41 (FI6 VH) và SEQ ID NO: 42 (chuỗi nhẹ biến đổi FI6). Các trình tự này và các dấu hiệu của chúng được kết hợp bằng cách viện dẫn.

AAV có thể được tạo ra từ các plasmit con thoi adhesin miễn dịch được mô tả ở trên sử dụng các kỹ thuật đã biết đối với người có hiểu biết trung bình trong lĩnh vực.

Các plasmit con thoi ví dụ khác như sau.

Trình tự của plasmit pN512_ACE FI6v3kg1-1A6 MAB_p3184 chứa chuỗi nhẹ dòng mầm kappa mà khác loại với nguồn của cả hai chuỗi nặng, 1A6 và FI6v3 được nêu trong SEQ ID NO: 14. Các trình tự mã hóa được dịch mã được nêu trong SEQ ID NO: 15 (chuỗi nhẹ hằng định), SEQ ID NO: 16 (chuỗi nặng biến đổi FI6), SEQ ID NO: 17 (CH1), và SEQ ID NO: 18 (CH2-3).

Các trình tự của vector trung gian mà mang các globulin miễn dịch chuỗi nặng và chuỗi nhẹ TCN032 được nêu ở SEQ ID NO: 30. Các trình tự axit amin được dịch mã được mã hóa bởi plasmit này bao gồm chuỗi nặng TCN032 nêu ở SEQ ID NO: 31; trình tự CH1 nêu ở SEQ ID NO: 32; chuỗi FI6 VH nêu ở SEQ ID NO: 33; trình tự CH1 nêu ở SEQ ID NO: 34 và trình tự CH2-3 nêu ở SEQ ID NO: 35.

Trình tự của plasmit mang các chuỗi nặng TCN032 và FI6 và hai kháng thể đồng biểu hiện có các tính đặc hiệu này được nêu trong SEQ ID NO: 43. Axit amin được dịch mã của chuỗi nặng biến đổi TCN032 nêu ở SEQ ID NO: 44, CH1 nêu ở SEQ ID NO: 45, đầu nối-CH2'-CH3' nêu ở SEQ ID NO: 46, FI6 VH nêu ở SEQ ID NO: 47, CH1 nêu ở SEQ ID NO: 48, CH2-3 nêu ở SEQ ID NO: 49, và gen kháng ampicillin nêu ở SEQ ID NO: 50. Các trình tự này và các đặc điểm của chúng được kết hợp vào đây bằng cách viện dẫn.

(Phần văn bản tự do của danh mục trình tự)

Thông tin sau đây được đề xuất cho trình tự chứa phân ký tự tự do theo ký hiệu số <223>.

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
1	<223> Trình tự tổng hợp mã hóa chuỗi nặng FI6 <220> <221> CDS <222> (1)..(705) <223> hằng định FI6
3	<223> plasmit mang kháng thể FI6 và 1A6 <220> <221> tín hiệu_polyA <222> (191)..(239) <223> tổng hợp_polyA <220> <221> đặc điểm_đa dạng <222> (246)..(914) <223> bổ sung - CH'2-3 <220> <221> đặc điểm_đa dạng <222> (915)..(1235) <223> bổ sung - CH'1 <220> <221> đặc điểm_đa dạng <222> (1236)..(1598) <223> bổ sung - 1A6\ VH <220> <221> đặc điểm_đa dạng <222> (1599)..(1655) <223> bổ sung - đoạn dẫn <220> <221> đặc điểm_đa dạng <222> (1734)..(2202)

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<223> Vùng tăng cường <220> <221>đặc điểm_đa dạng <222> (2388)..(2444) <223> đoạn dẫn <220> <221> CDS <222> (2445)..(2777) <223> FI6\VL <220> <221>đặc điểm_đa dạng <222> (3183)..(3242) <223> đoạn dẫn <220> <221> CDS <222> (3243)..(3629) <223> FI6\XH <220> <221> CDS <222> (3630)..(3950) <223> CH1 <220> <221> CDS <222> (3951)..(4619) <223> CH2-3 <220> <221>tín hiệu_polyA <222> (4626)..(4703) <223>TKpA ngắn <220> <221>đặc điểm_đa dạng <222> (6995)..(7283) <223> COL\E1\Origin

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
8	<223> Plasmit mã hóa các kháng thể đơn dòng FI6 và C05 <220> <221>tín hiệu_polyA <222> (204)..(252) <223> tổng hợp\polyA <220> <221>đặc điểm_đa dạng <222> (259)..(927) <223> bổ sung - CH'2-3 <220> <221>đặc điểm_đa dạng <222> (928)..(1248) <223> bổ sung - CH'1 <220> <221>đặc điểm_đa dạng <222> (1251)..(1668) <223> bổ sung - C05\ VH <220> <221>đặc điểm_đa dạng <222> (1669)..(1719) <223> bổ sung - đoạn dẫn <220> <221>đặc điểm_đa dạng <222> (1729)..(1979) <223> bổ sung - CMV\mp2 <220> <221>đặc điểm_đa dạng <222> (1798)..(2266) <223> Vùng tăng cường <220> <221>đặc điểm_đa dạng

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<222> (2267)..(2392) <223> CMV\mp2 <220> <221> CDS <222> (2509)..(2841) <223> FI6\VL <220> <221> CDS <222> (2842)..(3162) <223> CL <220> <221>đặc điểm đa dạng <222> (3247)..(3306) <223> đoạn dẫn <220> <221> CDS <222> (3307)..(3693) <223> FI6\XH <220> <221> CDS <222> (3694)..(4014) <223> CH1 <220> <221> CDS <222> (4015)..(4683) <223> CH2-3 <220> <221>polyA_ tín hiệu <222> (4690)..(4767) <223>TKpA ngắn
14	<223> Plasmid mã hóa các kháng thể đơn dòng FI6 và 1A6 tổng hợp

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<220> <221>tín hiệu_polyA <222> (191)..(239) <223> tổng hợp\polyA <220> <221>đặc điểm_đa dạng <222> (246)..(914) <223> bổ sung - CH'2-3 <220> <221>đặc điểm_đa dạng <222> (915)..(1235) <223> bổ sung - CH'1 <220> <221>đặc điểm_đa dạng <222> (1236)..(1598) <223> bổ sung - 1A6\ VH <220> <221>đặc điểm_đa dạng <222> (1599)..(1655) <223> bổ sung - đoạn dẫn <220> <221>đặc điểm_đa dạng <222> (1665)..(1733) <223> bổ sung - CMV\mp2 <220> <221>đặc điểm_đa dạng <222> (1732)..(2202) <223> Vùng tăng cường <220> <221>đặc điểm_đa dạng <222> (2203)..(2328) <223> CMV\mp1 <220>

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<221>đặc điểm_đa dạng <222> (2388)..(2444) <223> đoạn dẫn <220> <221>đặc điểm_đa dạng <222> (2445)..(2789) <223> KGL <220> <221> CDS <222> (2784)..(3104) <223> CL <220> <221>đặc điểm_đa dạng <222> (3189)..(3248) <223> đoạn dẫn <220> <221> CDS <222> (3249)..(3635) <223> FI6\ VH <220> <221> CDS <222> (3636)..(3956) <223> CH1 <220> <221> CDS <222> (3957)..(4625) <223> CH2-3 <220> <221>tín hiệu_polyA <222> (4632)..(4709) <223>TKpA ngắn
19	<223> Plasmit mang kháng thể đơn dòng FI6 và CR8033

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<220> <221>tín hiệu_polyA <222> (173)..(221) <223> tổng hợp\polyA <220> <221>đặc điểm_đa dạng <222> (228)..(896) <223> bổ sung - CH'2-3 <220> <221>đặc điểm_đa dạng <222> (897)..(1217) <223> bổ sung - CH'1 <220> <221>đặc điểm_đa dạng <222> (1218)..(1604) <223> bổ sung - CR8033\VH <220> <221>đặc điểm_đa dạng <222> (1605)..(1655) <223> bổ sung - đoạn dẫn <220> <221>đặc điểm_đa dạng <222> (1665)..(1733) <223> bổ sung - CMV\mp2 <220> <221>đặc điểm_đa dạng <222> (1734)..(2202) <223> Vùng tăng cường <220> <221>đặc điểm_đa dạng <222> (2203)..(2328) <223> CMV\mp1

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<220> <221>đặc điểm_đa dạng <222> (2445)..(2789) <223> KGL <220> <221> CDS <222> (2784)..(3104) <223> CL <220> <221>đặc điểm_đa dạng <222> (3189)..(3248) <223> đoạn dẫn <220> <221> CDS <222> (3249)..(3635) <223> FI6\ VH <220> <221> CDS <222> (3636)..(3956) <223> CH1 <220> <221> CDS <222> (3957)..(4625) <223> CH2-3 <220> <221>đặc điểm_đa dạng <222> (3968)..(3968) <223> A -> T <220> <221>tín hiệu_polyA <222> (4632)..(4709) <223>TKpA ngắn
24	<220>

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<223> Plasmit mang các kháng thể đơn dòng FI6 và CR8033 <220> <221>tín hiệu_polyA <222> (191)..(239) <223> tổng hợp polyA <220> <221>đặc điểm_đa dạng <222> (246)..(914) <223> bổ sung - CH'2-3 <220> <221>đặc điểm_đa dạng <222> (915)..(1235) <223> bổ sung - CH'1 <220> <221>đặc điểm_đa dạng <222> (1236)..(1622) <223> bổ sung - CR8033\VH <220> <221>đặc điểm_đa dạng <222> (1623)..(1673) <223> bổ sung - đoạn dẫn <220> <221>đặc điểm_đa dạng <222> (1683)..(1751) <223> CMV\mp2 <220> <221>đặc điểm_đa dạng <222> (1752)..(2220) <223> Vùng tăng cường <220>

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<221>đặc điểm đa dạng <222> (2221)..(2346) <223> CMV\mp1 <220> <221>đặc điểm đa dạng <222> (2406)..(2462) <223> đoạn dẫn <220> <221> CDS <222> (2463)..(2795) <223> FI6\VL <220> <221> CDS <222> (2796)..(3116) <223> CL <220> <221>đặc điểm đa dạng <222> (3201)..(3260) <223> đoạn dẫn <220> <221> CDS <222> (3261)..(3647) <223> FI6\XH <220> <221> CDS <222> (3648)..(3968) <223> CH1 <220> <221> CDS <222> (3969)..(4637) <223> CH2-3 <220> <221>đặc điểm đa dạng

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<222> (3980)..(3980) <223> A -> T <220> <221>tín hiệu_polyA <222> (4644)..(4721) <223>TKpA ngắn
30	<223>EcoRV <220> <221>tín hiệu_polyA <222> (201)..(252) <223> bổ sung - tổng hợp_polyA <220> <221>đặc điểm_đa dạng <222> (268)..(588) <223> bổ sung - CL <220> <221>đặc điểm_đa dạng <222> (589)..(909) <223> bổ sung - TCN032\VL <220> <221>tín hiệu_polyA <222> (910)..(966) <223> bổ sung - đoạn dẫn <220> <221>đặc điểm_đa dạng <222> (1026)..(1094) <223> bổ sung - CMV\mp2 <220> <221>đặc điểm_đa dạng <222> (1095)..(1563) <223> Vùng tăng cường <220>

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<221>đặc điểm_đa dạng <222> (1564)..(1689) <223> CMV\mp1 <220> <221>đặc điểm_đa dạng <222> (1749)..(1805) <223> đoạn dẫn <220> <221> CDS <222> (1806)..(2165) <223> TCN032\ VH <220> <221> CDS <222> (2166)..(2459) <223> CH1 <220> <221>đặc điểm_đa dạng <222> (2460)..(3152) <223> khớp nối-CH2'-CH3' <220> <221>đặc điểm_đa dạng <222> (3239)..(3296) <223> đoạn dẫn <220> <221> CDS <222> (3297)..(3683) <223> FI6\ VH <220> <221> CDS <222> (3684)..(4004) <223> CH1 <220> <221> CDS

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<222> (4005)..(4673) <223> CH2-3 <220> <221>tín hiệu_polyA <222> (4693)..(4770) <223>TKpA ngắn
36	<223> adhesin miễn dịch FI6 và CO5 <220> <221>tín hiệu_polyA <222> (201)..(432) <223> bổ sung - SV40\polyA <220> <221>đặc điểm_đa dạng <222> (453)..(1121) <223> bổ sung - CH'2-3 <220> <221>đặc điểm_đa dạng <222> (1125)..(1457) <223> bổ sung - C05\VL <220> <221>đặc điểm_đa dạng <222> (1458)..(1502) <223> SL\from\3bn201co <220> <221>đặc điểm_đa dạng <222> (1503)..(1916) <223> bổ sung - C05\XH <220> <221>đặc điểm_đa dạng <222> (1965)..(1973) <223> đoạn dẫn <220>

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<221>đặc điểm_đa dạng <222> (2371)..(2412) <223> bổ sung - CMV\mp2 <220> <221>đặc điểm_đa dạng <222> (2413)..(2881) <223> vùng tăng cường <220> <221>đặc điểm_đa dạng <222> (2882)..(3007) <223> CMV\mp1 <220> <221>đặc điểm_đa dạng <222> (3067)..(3055) <223> đoạn dẫn <220> <221> CDS <222> (3124)..(3510) <223> FI6\ VH <220> <221>đặc điểm_đa dạng <222> (3511)..(3555) <223> SL\from\3bn201co <220> <221> CDS <222> (3556)..(3888) <223> FI6\VL <220> <221> CDS <222> (3892)..(4560) <223> CH2-3 <220> <221>tín hiệu_polyA

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<222> (4581)..(4812) <223> SV40\polyA
40	<223> các adhesin miễn dịch FI6 và CR8033 <220> <221>tín hiệu_polyA <222> (201)..(432) <223> bổ sung - SV40\polyA <220> <221>đặc điểm_đa dạng <222> (453)..(1121) <223> bổ sung - CH'2-3 <220> <221>đặc điểm_đa dạng <222> (1125)..(1460) <223> bổ sung - 033\VL <220> <221>đặc điểm_đa dạng <222> (1461)..(1505) <223> SL\from\3bn201co <220> <221>đặc điểm_đa dạng <222> (1506)..(1886) <223> bổ sung - 033\VH <220> <221>đặc điểm_đa dạng <222> (1935)..(1946) <223> bổ sung - đoạn dẫn <220> <221>đặc điểm_đa dạng <222> (2341)..(2382) <223> bổ sung - CMV\mp2 <220>

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<221>đặc điểm_đa dạng <222> (2383)..(2851) <223> vùng tăng cường <220> <221>đặc điểm_đa dạng <222> (2852)..(2977) <223> CMV\mp1 <220> <221>đặc điểm_đa dạng <222> (3073)..(3045) <223> đoạn dẫn <220> <221> CDS <222> (3094)..(3480) <223> FI6\ VH <220> <221>đặc điểm_đa dạng <222> (3481)..(3525) <223> SL\from\3bn201co <220> <221> CDS <222> (3526)..(3858) <223> FI6\VL <220> <221>đặc điểm_đa dạng <222> (3862)..(4530) <223> CH2-3 <220> <221>tín hiệu_polyA <222> (4551)..(4782) <223> SV40\polyA
43	<223> Plasmit mang các kháng thể đơn dòng TCN032 và Fi6

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<220> <221>lặp_vùng <222> (14)..(143) <220> <221>tín hiệu_polyA <222> (204)..(252) <223> tổng hợp polyA <220> <221>đặc điểm_đa dạng <222> (261)..(267) <223> cat-xet chặn (bổ sung) <220> <221>đặc điểm_đa dạng <222> (268)..(588) <223> chuỗi nhẹ hằng định (trên chuỗi bổ sung) <220> <221>đặc điểm_đa dạng <222> (967)..(971) <223>Kozak (nằm trên chuỗi bổ sung) <220> <221>đặc điểm_đa dạng <222> (972)..(1019) <223> c-myc 5' UTR (nằm trên chuỗi bổ sung) <220> <221>đặc điểm_đa dạng <222> (1026)..(1094) <223> CMV\mp2 <220> <221> vùng tăng cường <222> (1026)..(1094)

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<220> <221>đặc điểm_đa dạng <222> (1564)..(1689) <220> <221>đặc điểm_đa dạng <222> (1696)..(1743) <223> c-myc 5' UTR <220> <221>đặc điểm_đa dạng <222> (1744)..(1748) <223>Kozak <220> <221>đặc điểm_đa dạng <222> (1749)..(1805) <223> đoạn dẫn <220> <221> CDS <222> (1806)..(2165) <223> chuỗi nặng biến đổi TCN032 <220> <221>vùng_lập <222> (1845)..(4974) <223> đoạn cuối lập đảo ngược <220> <221>vùng_lập <222> (1845)..(4974) <223> đoạn lập đầu cuối đảo ngược (nằm trên chuỗi bổ sung) <220> <221> CDS

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<222> (2166)..(2459) <223> CH1 <220> <221>misc <222> (2166)..(2459) <223> CH1 <220> <221> CDS <222> (2460)..(3152) <223> khớp nối-CH2'-CH3' <220> <221>đặc điểm_đa dạng <222> (3153)..(3164) <223> vị trí phân cắt furin <220> <221>đặc điểm_đa dạng <222> (3165)..(3236) <223> Liên kết F2 <220> <221>đặc điểm_đa dạng <222> (3239)..(3296) <220> <221>đặc điểm_đa dạng <222> (3239)..(3296) <220> <221> CDS <222> (3297)..(3683) <223> FI6 VH <220> <221> CDS <222> (3684)..(4004) <223> CH1

SEQ ID NO: (chứa ký tự tự do)	Ký tự tự do theo <223>
	<220> <221> CDS <222> (4005)..(4673) <223> CH2-3 <220> <221>đặc điểm_đa dạng <222> (4674)..(4680) <223> Cat-xet chặn <220> <221>đặc điểm_đa dạng <222> (4674)..(4680) <220> <221>tín hiệu_polyA <222> (4693)..(4770) <223>TKpA ngắn <220> <221>rep_gốc <222> (5151)..(5606) <220> <221> CDS <222> (5737)..(6594) <223> Amp-R <220> <221>đặc điểm_đa dạng <222> (6768)..(7356) <223> col\E1\origin

Sáng chế chứa các trình tự và danh mục trình tự, được kết hợp vào đây bằng cách viện dẫn. Toàn bộ công bố đơn, sáng chế, và đơn sáng chế được liệt kê ở sáng chế, và đơn sáng chế tạm thời Mỹ số 61/992,649, được nộp ngày 13/05/2014, được yêu cầu hưởng quyền ưu tiên, được kết hợp vào đây bằng cách viện dẫn toàn bộ như thể mỗi công bố riêng hoặc đơn sáng chế được nêu cụ thể và

riêng biệt để được kết hợp bằng cách viện dẫn. Mặc dù sáng chế trên được mô tả khá chi tiết bằng cách minh họa và ví dụ nhằm mục đích hiểu rõ, hiển nhiên với người có hiểu biết trung bình trong lĩnh vực là theo các hướng dẫn của sáng chế rằng một số thay đổi và cải biến có thể được thực hiện theo đó mà không vượt ra khỏi phạm vi hoặc tinh thần của yêu cầu bảo hộ kèm theo.

Yêu cầu bảo hộ

1. Virut liên quan đến adeno tái tổ hợp (AAV) có vỏ protein của virut AAV và được bọc bên trong đó axit nucleic khác loại mà biểu hiện hai kháng thể đơn đặc hiệu chức năng trong tế bào, trong đó AAV tái tổ hợp này chứa:

đoạn lặp đầu cuối đảo ngược (ITR) của AAV tại đầu 5’;

cat-xet biểu hiện thứ nhất mà mã hóa ít nhất khung đọc mở thứ nhất (ORF) cho globulin miễn dịch thứ nhất dưới sự điều khiển của các trình tự điều khiển điều hòa mà định hướng sự biểu hiện của nó;

cat-xet biểu hiện thứ hai mà chứa ORF thứ hai, liên kết, và ORF thứ ba dưới sự điều khiển của các trình tự điều khiển điều hòa mà định hướng biểu hiện của chúng, trong đó ORF thứ hai và thứ ba là dùng cho cấu trúc globulin miễn dịch thứ hai và thứ ba; và

ITR của AAV tại đầu 3’,

trong đó chuỗi nhẹ globulin miễn dịch, chuỗi nặng globulin miễn dịch thứ nhất và chuỗi nặng globulin miễn dịch thứ hai được mã hóa.

2. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó AAV tái tổ hợp này còn biểu hiện kháng thể đặc hiệu kép.

3. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó AAV tái tổ hợp này chứa vùng tăng cường hai hướng nằm giữa cat-xet biểu hiện thứ nhất và cat-xet biểu hiện thứ hai.

4. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó ORF thứ nhất mã hóa chuỗi nhẹ globulin miễn dịch, ORF thứ hai mã hóa chuỗi nặng globulin miễn dịch thứ nhất và ORF thứ ba mã hóa chuỗi nặng thứ hai, nhờ đó các cấu trúc kháng thể chức năng được biểu hiện có hai chuỗi nặng khác nhau có các tính đặc hiệu khác nhau mà có chung chuỗi nhẹ.

5. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó ít nhất một trong số ORF thứ hai và thứ ba chứa các trình tự mã hóa Fc biến đổi.
6. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó liên kết trong cat-xet thứ hai chứa liên kết được chọn từ IRES hoặc F2A.
7. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó các trình tự điều khiển điều hòa đối với cat-xet biểu hiện thứ nhất và/hoặc cat-xet biểu hiện thứ hai chứa vùng khởi đầu tối thiểu.
8. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó các trình tự điều khiển điều hòa đối với cat-xet biểu hiện thứ nhất và/hoặc cat-xet biểu hiện thứ hai chứa polyA tối thiểu hoặc tổng hợp.
9. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó cat-xet biểu hiện thứ nhất là xistron kép và chứa ORF nữa.
10. Virut liên quan đến adeno tái tổ hợp theo điểm 9, trong đó mỗi trong số ORF chứa scFv.
11. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó vector chứa polyA hai hướng nằm giữa cat-xet biểu hiện thứ nhất và cat-xet biểu hiện thứ hai.
12. Virut liên quan đến adeno tái tổ hợp theo điểm 10, trong đó cat-xet biểu hiện thứ nhất chứa vùng tăng cường và vùng khởi đầu tối thiểu.
13. Virut liên quan đến adeno tái tổ hợp theo điểm 12, trong đó cat-xet biểu hiện thứ hai chứa vùng tăng cường và vùng khởi đầu tối thiểu.
14. Virut liên quan đến adeno tái tổ hợp theo điểm 8, trong đó cat-xet biểu hiện thứ nhất và thứ hai cùng nhau biểu hiện hai Fab.
15. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó hai kháng thể đơn đặc hiệu có các tính đặc hiệu khác nhau.

16. Virut liên quan đến adeno tái tổ hợp theo điểm 1, trong đó hai kháng thể đơn đặc hiệu độc lập được chọn từ kháng thể đơn dòng, adhesin miễn dịch, Fab, kháng thể hai chức năng, và các tổ hợp của chúng.

17. Virut liên quan đến adeno (AAV) tái tổ hợp có vỏ protein AAV và được đóng gói trong đó axit nucleic khác loại mà biểu hiện hai kháng thể đơn đặc hiệu chức năng, trong đó AAV tái tổ hợp biểu hiện kháng thể đơn dòng thứ nhất có tính đặc hiệu thứ nhất, kháng thể đơn dòng thứ hai có tính đặc hiệu khác với kháng thể đơn dòng thứ nhất, và kháng thể đặc hiệu kép, trong đó AAV tái tổ hợp bao gồm:

đoạn lặp đầu cuối đảo ngược (ITR) của AAV tại đầu 5’;

cat-xet biểu hiện thứ nhất mà mã hóa ít nhất khung đọc mở thứ nhất (ORF) đối với globulin miễn dịch thứ nhất dưới sự điều khiển của các trình tự điều khiển điều hòa mà định hướng sự biểu hiện của nó;

cat-xet biểu hiện thứ hai mà chứa ORF thứ hai, liên kết, và ORF thứ ba dưới sự điều khiển của các trình tự điều khiển điều hòa mà định hướng sự biểu hiện của nó, trong đó ORF thứ hai và thứ ba là dùng cho cấu trúc globulin miễn dịch thứ hai và thứ ba; và

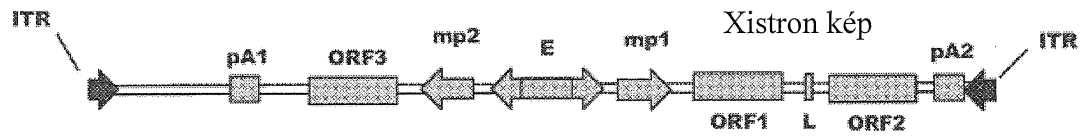
ITR của AAV tại đầu 3’,

trong đó chuỗi nhẹ globulin miễn dịch, chuỗi nặng globulin miễn dịch thứ nhất và chuỗi nặng globulin miễn dịch thứ hai được mã hóa.

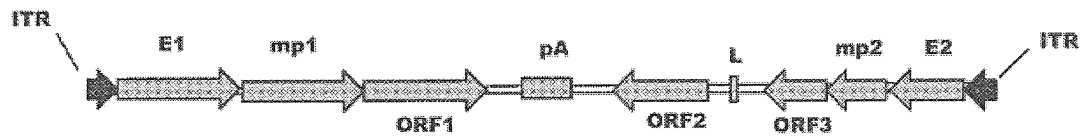
18. Dược phẩm chứa virut liên quan đến adeno tái tổ hợp theo điểm bất kỳ trong số các điểm từ 1 đến 17 và chất mang dược dụng.

19. Chế phẩm chứa hai kháng thể đơn đặc hiệu chức năng có các tính đặc hiệu khác nhau, trong đó mỗi kháng thể trong số các kháng thể có cùng chuỗi nhẹ và chuỗi nặng khác nhau, trong đó chuỗi nhẹ là từ nguồn khác loại với nguồn của chuỗi nặng đối với một hoặc cả hai kháng thể, và trong đó các kháng thể được sản sinh bằng cách biểu hiện từ AAV tái tổ hợp theo điểm bất kỳ trong số các điểm từ 1 đến 17.

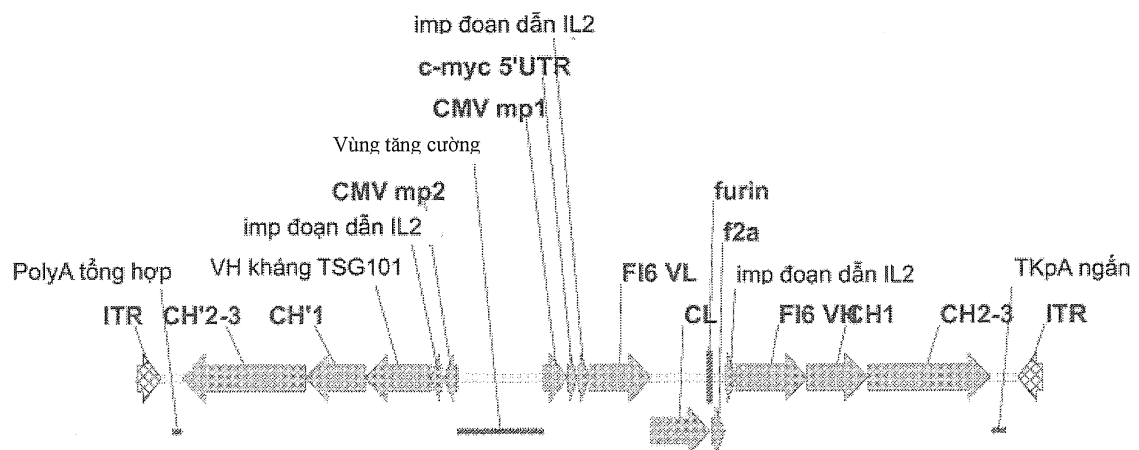
20. Virut liên quan đến adeno tái tổ hợp theo điểm 17, trong đó ORF thứ nhất mã hóa chuỗi nhẹ globulin miễn dịch, ORF thứ hai mã hóa chuỗi nặng globulin miễn dịch thứ nhất và ORF thứ ba mã hóa chuỗi nặng thứ hai, trong đó các cấu trúc kháng thể chức năng được biểu hiện có hai chuỗi nặng khác nhau với các tính đặc hiệu khác nhau mà có chung chuỗi nhẹ.



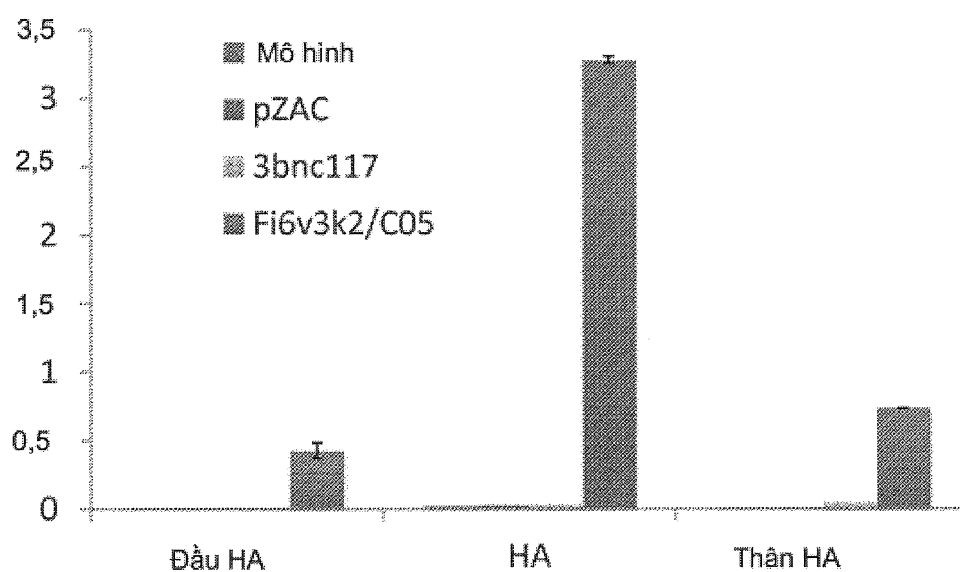
Hình 1A



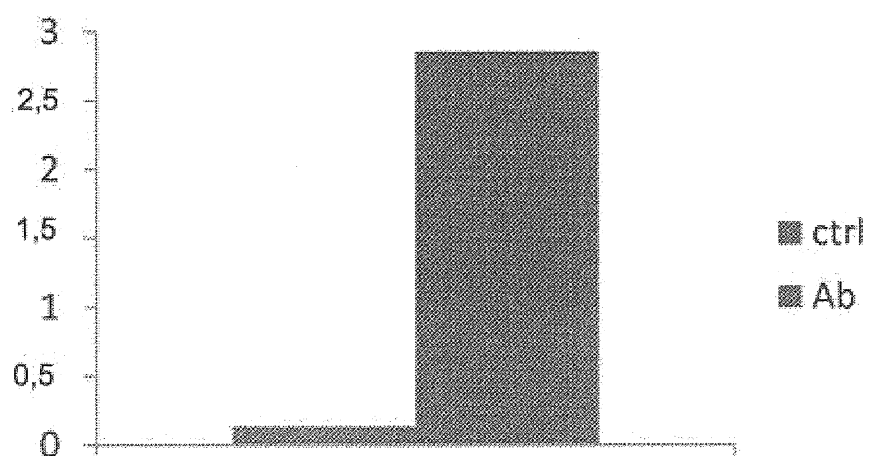
Hình 1B



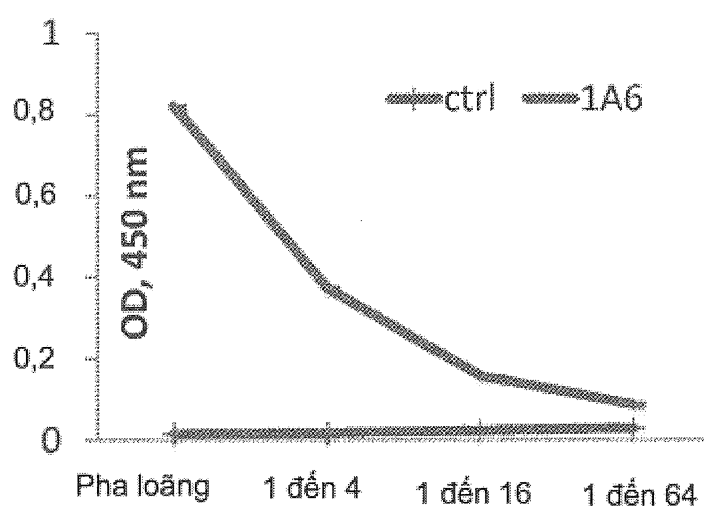
Hình 2



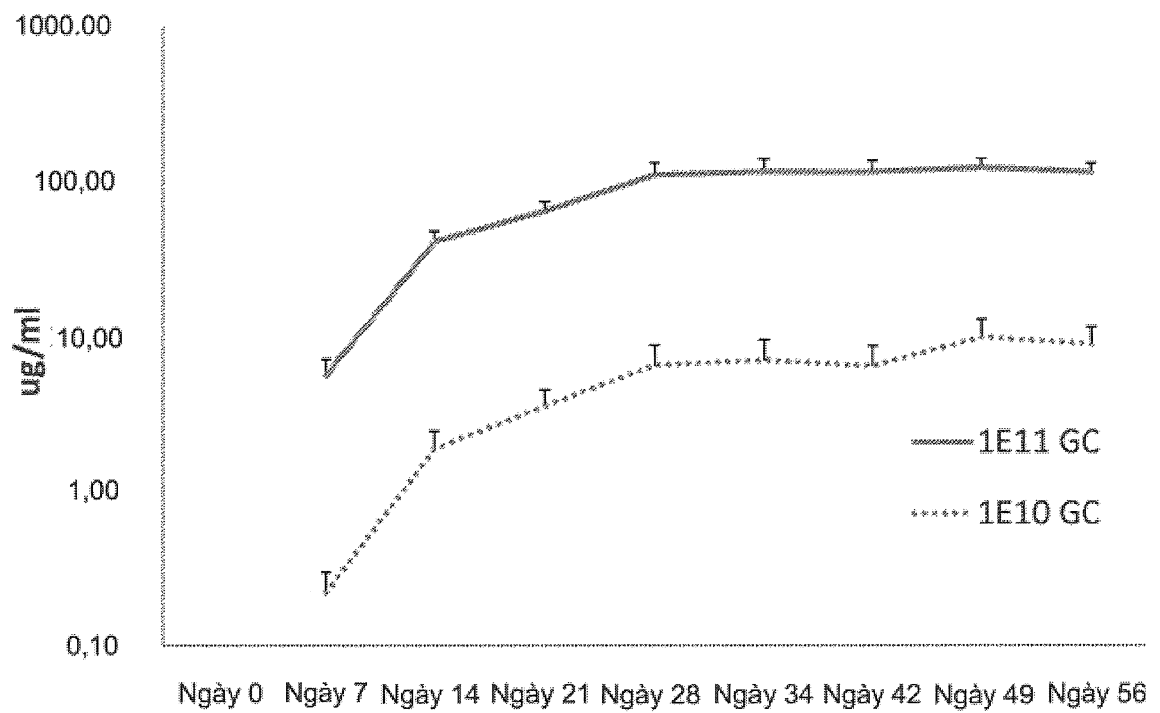
Hình 3



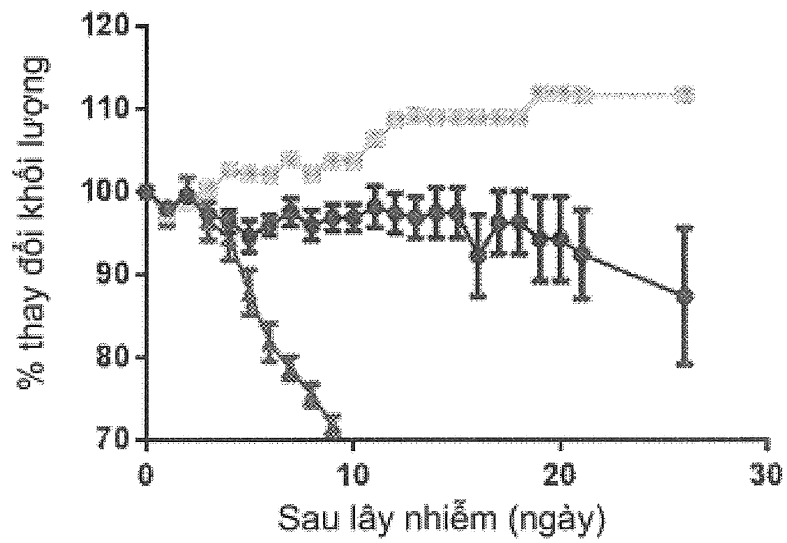
Hình 4A



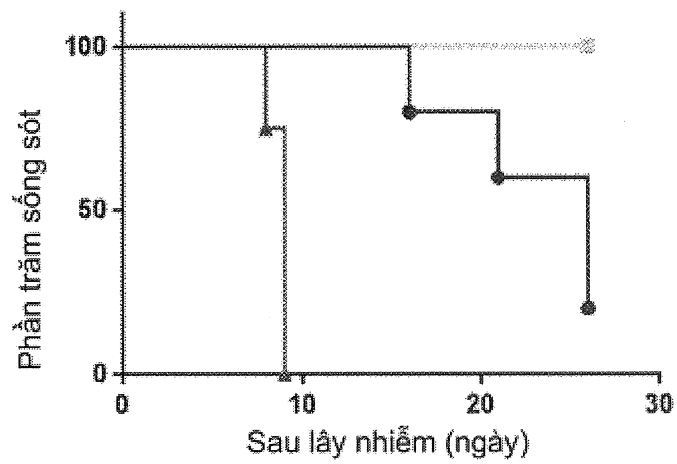
Hình 4B



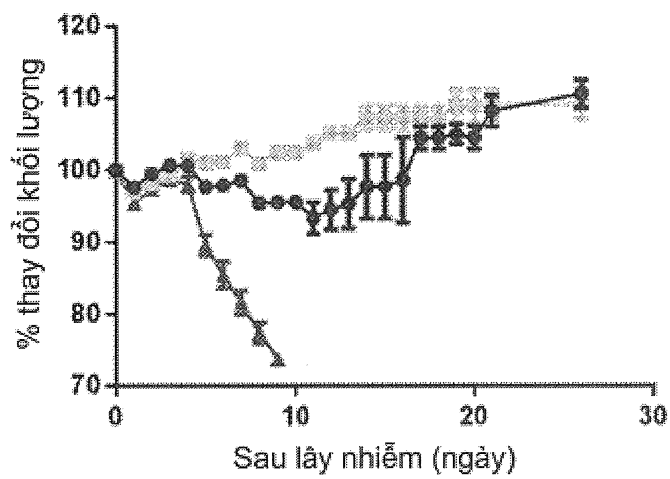
Hình 5



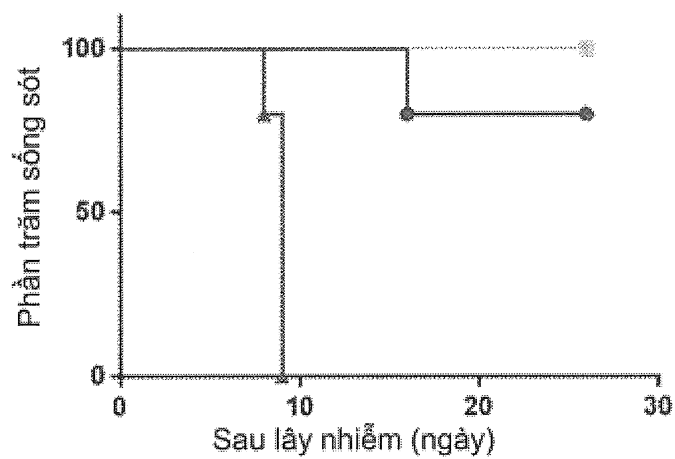
Hình 6A



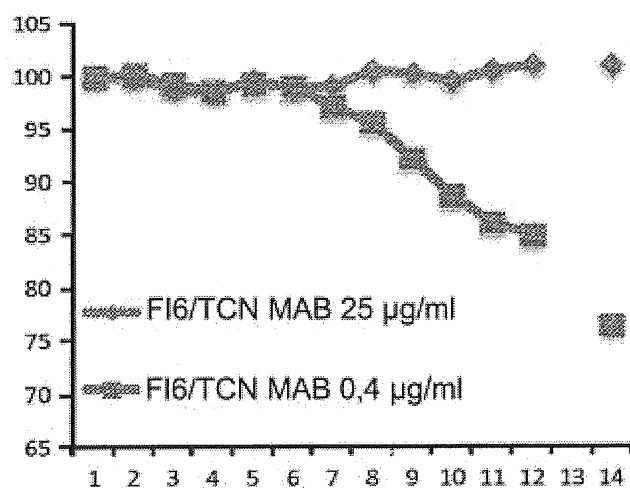
Hình 6B



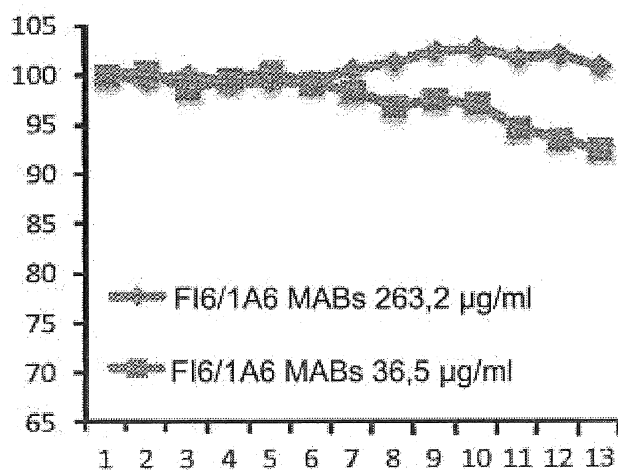
Hình 7A



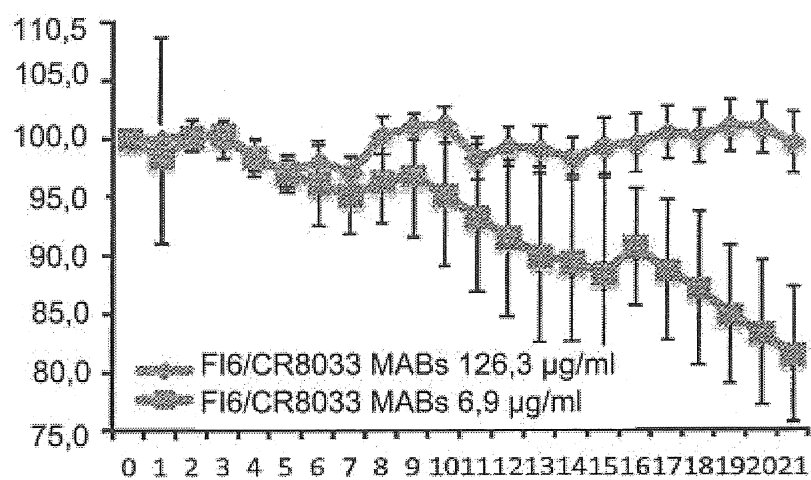
Hình 7B



Hình 8A



Hình 8B



Hình 8C

DANH MỤC TRÌNH TỰ

<110> The Trustees of the University of Pennsylvania
 <120> Virut liên quan đến adeno (AAV) tái tổ hợp, chế phẩm và dược phẩm chứa virut liên quan đến adeno tái tổ hợp này
 <130> UPN-14-7032PCT
 <150> 61/992649
 <151> 2014-05-13
 <160> 50
 <170> PatentIn phiên bản 3.5
 <210> 1
 <211> 705
 <212> ADN
 <213> Trình tự nhân tạo
 <220>
 <223> Trình tự tổng hợp mã hóa chuỗi nặng FI6
 <220>
 <221> CDS
 <222> (1)..(705)
 <223> Vùng hằng định FI6
 <400> 1
 gcg gcg cct aag agc tgc gac aag acc cac acc tgt ccc ccc tgc cct 48
 Ala Ala Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
 1 5 10 15
 gcc cct gaa ctg ctg gga ggc ccc agc gtg ttc ctg ttc ccc cca aag 96
 Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
 20 25 30
 ccc aag gac acc ctg atg atc agc cgg acc ccc gaa gtg acc tgc gtg 144
 Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
 35 40 45
 gtg gtg gac gtg tcc cac gag gac cct gaa gtg aag ttc aat tgg tac 192
 Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
 50 55 60
 gtg gac ggc gtg gaa gtg cac aac gcc aag acc aag ccc aga gag gaa 240
 Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu
 65 70 75 80
 cag tac aac agc acc tac cgg gtg gtg tcc gtg ctg acc gtg ctg cac 288
 Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
 85 90 95

cag gac tgg ctg aac ggc aaa gag tac aag tgc aag gtg tcc aac aag 336
 Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
 100 105 110

gcc ctg cct gcc ccc atc gag aaa acc atc agc aag gcc aag ggc cag 384
 Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
 115 120 125

ccc cgc gag cct cag gtg tgc aca ctg ccc ccc agc cgg gaa gag atg 432
 Pro Arg Glu Pro Gln Val Cys Thr Leu Pro Pro Ser Arg Glu Glu Met
 130 135 140

acc aag aac cag gtg tcc ctg acc tgc ctg gtc aag ggc ttc tac ccc 480
 Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
 145 150 155 160

agc gat atc gcc gtg gaa tgg gag agc aac ggc cag ccc gag aac aac 528
 Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
 165 170 175

tac aag acc acc ccc cct gtg ctg gac agc gac ggc tca ttc ttc ctg 576
 Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
 180 185 190

tac agc aag ctg acc gtg gac aag agc cgg tgg cag cag ggc aac gtg 624
 Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
 195 200 205

ttc agc tgc agc gtg atg cac gag gcc ctg cac aac cac tac acc cag 672
 Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
 210 215 220

aag tcc ctg agc ctg agc ccc ggc aag tga tga 705
 Lys Ser Leu Ser Leu Ser Pro Gly Lys
 225 230

<210> 2
 <211> 233
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 2

Ala Ala Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
 1 5 10 15

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
 20 25 30

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val
 35 40 45

Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr
 50 55 60

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

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
 85 90 95

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
 100 105 110

Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
 115 120 125

Pro Arg Glu Pro Gln Val Cys Thr Leu Pro Pro Ser Arg Glu Glu Met
 130 135 140

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro
 145 150 155 160

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
 165 170 175

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
 180 185 190

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
 195 200 205

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
 210 215 220

Lys Ser Leu Ser Leu Ser Pro Gly Lys
 225 230

<210> 3
 <211> 7722
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> plasmit mang các kháng thể FI6 và 1A6

<220>
 <221> tín hiệu polyA
 <222> (191)..(239)
 <223> tổng hợp\polyA

<220>
 <221> đặc điểm đa dạng
 <222> (246)..(914)
 <223> bổ sung- CH'2-3

<220>
 <221> đặc điểm đa dạng
 <222> (915)..(1235)
 <223> bổ sung- CH'1

<220>
 <221> đặc điểm đa dạng
 <222> (1236)..(1598)
 <223> bổ sung- 1A6\VH

<220>
 <221> đặc điểm đa dạng
 <222> (1599)..(1655)
 <223> bổ sung- đoạn dẫn

<220>
 <221> đặc điểm đa dạng
 <222> (1734)..(2202)
 <223> Đoạn tăng cường

<220>
 <221> đặc điểm đa dạng
 <222> (2388)..(2444)
 <223> đoạn dẫn

<220>
 <221> CDS
 <222> (2445)..(2777)
 <223> FI6\VL

<220>
 <221> đặc điểm đa dạng
 <222> (3183)..(3242)
 <223> đoạn dẫn

<220>
 <221> CDS
 <222> (3243)..(3629)
 <223> FI6\XH

<220>
 <221> CDS
 <222> (3630)..(3950)
 <223> CH1

<220>
 <221> CDS
 <222> (3951)..(4619)
 <223> CH2-3

<220>
 <221> tín hiệu_polyA
 <222> (4626)..(4703)
 <223> TKpA ngắn

<220>
 <221> đặc điểm_đa dạng
 <222> (6995)..(7283)
 <223> COL\E1\gốc

<400> 3
 ctgcgcgctc gctcgcctcac tgaggccgcc cgggcaaagc ccgggcgctcg ggcgaccttt 60
 ggtcgcccg gctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact 120
 aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct 180
 aggaagatct cacacaaaa accaacacac agatgtaatg aaaataaaga tattttatatt 240
 tatcacttcc cgggggctcag gctcagggac ttctgggtgt agtgggtgtg cagggcctcg 300
 tgcatacagc tgcagctgaa cacgttgccc tgctgccacc ggctcttgtc cacggtcagc 360
 ttgctataca ggaagaatga gccgtcgctg tccagcacag ggggggtggg cttgtagttg 420
 ttctcgggct ggccgttgct ctccatttcc acggcgatct cgctggggta gaagcccttg 480
 accaggcagg tcagggacac ctggttcttg gtcattcttt cccggctggg gggcagtggt 540
 tagacctgag gctcgcgggg ctggcccttg gccttgctga tggttttctc gatgggggca 600
 ggcagggcct tgttgagacac cttgcacttg tactctttgc cgttcagcca gtccgtgtgc 660
 agcacggtca gcacggacac caccggttag gtgctgttgt actgttcttc tctgggcttg 720
 gtcttgggct tgtgcacttc cacgccgtcc acgtaccaat tgaacttcac ttcagggtcc 780
 tcgtgggaca cgtccaccac cacgcaggtc acttcggggg tccggctgat catcagggtg 840

tccttgggct ttggggggaa caggaacacg ctggggcctc ccagcagttc aggggcaggg	900
caggggggac acgtgtgggt cttgtgcgag ctcttaggtt ccacccgctt gtccaccttg	960
gtgttgctgg gcttgtgggt cacgttgacg atgtaggtct gggtgcccag gctgctgctg	1020
ggcacggtga ccacgctgct caggctatac aggcgctgc tctgcagcac ggctggaaag	1080
gtgtgcacgc cgctggtcag ggcgccagag ttccaggaca cggtcacggg ctcggggaag	1140
tagtccttga ccaggcagcc cagggcggct gttccgccag aggtgctctt gctgctaggg	1200
gccagagggg acacgcttgg tcccttggtg ctggcgctcg agacggtcac cagggttccc	1260
tgtccccagt aatccattcc tccgctggcg attccgctcc gatccttggc gcagtagtac	1320
acggcggtat cctcggcccc caggctgttc atctgcaggt acaggggtgtt cttgctgttg	1380
gcccggctga tgggtgaaccg tcccttcacg ctatcggcgt agtacttggt gtttccatcg	1440
tagctgatca cggccaccca ctccagtccc tttcctgggg cctgccgcac ccagtgcatt	1500
ccgtaatcgc tgaaggtgaa tccgctggcg gcgcagctca gccgcaggct ccgtcctggc	1560
tgcaccactc ctctccgct ctctgcagc tgcacctgtg aattcgtcac cagggccagg	1620
ctcagggcga tcagcagcag cagctgcatg cgcattggtg cggcgcgatc tgacggttca	1680
ctaaacgagc tctgcttata taggcctccc accgtacacg ccacctcgac atacctagtt	1740
attaatagta atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgtta	1800
cataacttac ggtaaatggc ccgcctggct gaccgcccac cgacccccgc ccattgacgt	1860
caataatgac gtatgttccc atagtaacgc caatagggac tttccattga cgtcaatggg	1920
tggagtattht acggtaaaact gccacttgg cagtacatca agtgtatcat atgccaagta	1980
cgccccctat tgacgtcaat gacggtaaht ggcccgcctg gcattatgcc cagtacatga	2040
ccttatggga ctttctact tggcagtaca tctacgtatt agtcatcgct attaccatgg	2100
tgatgcgggtt ttggcagtac atcaatgggc gtggatagcg gtttgactca cggggatttc	2160
caagtctcca cccattgac gtcaatggga gtttgttttg gcaccaaahat caacgggact	2220
ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat gggcggtagg cgtgtacggc	2280
gggaggtcta tataagcaga gctggtttag tgaaccgtca gatccgctgg gcactttgca	2340
ctggaactta caacacccga gcaaggacgc gactctgccg cccacccatg cgcattgcagc	2400
tgctgctgct gatcgccctg agcctggccc tggtagcaaa cagc gat atc gtc atg	2456
Asp Ile Val Met	

1

acc cag agc cca gat agc ctg gcc gtg agc ctg gga gag cgg gcc acc	2504
Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly Glu Arg Ala Thr	
5 10 15 20	
atc aac tgc aag agc agc cag agc gtg acc ttc aac tac aag aac tac	2552
Ile Asn Cys Lys Ser Ser Gln Ser Val Thr Phe Asn Tyr Lys Asn Tyr	
25 30 35	
ctg gcc tgg tac cag cag aag cca gga cag cca cca aag ctg ctg atc	2600
Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile	
40 45 50	
tac tgg gcc agc acc cgg gag agc gga gtg cca gat cgg ttc agc gga	2648
Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg Phe Ser Gly	
55 60 65	
agc gga agc gga acc gat ttc acc ctg acc atc agc agc ctg cag gcc	2696
Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Ala	
70 75 80	
gag gat gtg gcc gtg tac tac tgc cag cag cac tac cgg acc cca cca	2744
Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln His Tyr Arg Thr Pro Pro	
85 90 95 100	
acc ttc gga cag gga acc aag gtg gag atc aag cgtacggtgg cgcgcccaag	2797
Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys	
105 110	
cgtgttcac tc tccccaccaa gcgatgagca gctgaagagc ggaaccgcca gcgtggtgtg	2857
cctgctgaac aacttctacc cacgggagggc caaggtgcag tggaaggtgg ataacgcct	2917
gcagagcggg aacagccagg agagcgtgac cgagcaggat agcaaggata gcacctacag	2977
cctgagcagc accctgaccc tgagcaaggc cgattacgag aagcacaagg tgtacgcctg	3037
cgaggtgacc caccagggac tgagcagccc agtgaccaag agcttcaacc gcggagagtg	3097
ccggaagcgg cgggccccag tgaagcagac cctgaacttc gatctgctga agctggccgg	3157
agatgtggag agcaaccag gaccaatgta cagaatgcag ctgctgagct gcacgcct	3217
gagcctggcc ctggtgacca acagc cag gtg caa cta gtg gag agc gga gga	3269
Gln Val Gln Leu Val Glu Ser Gly Gly	
115 120	
gga gtg gtg cag cca gga cgg agc ctg cgg ctg agc tgc gcc gcc agc	3317
Gly Val Val Gln Pro Gly Arg Ser Leu Arg Leu Ser Cys Ala Ala Ser	
125 130 135	
gga ttc acc ttc agc acc tac gcc atg cac tgg gtg cgg cag gcc cca	3365
Gly Phe Thr Phe Ser Thr Tyr Ala Met His Trp Val Arg Gln Ala Pro	

140										145					150					
gga	aag	gga	ctg	gag	tgg	gtg	gcc	gtg	atc	agc	tac	gat	gcc	aac	tac	3413				
Gly	Lys	Gly	Leu	Glu	Trp	Val	Ala	Val	Ile	Ser	Tyr	Asp	Ala	Asn	Tyr					
		155					160					165								
aag	tac	tac	gcc	gat	agc	gtg	aag	gga	cgg	ttc	acc	atc	agc	cgg	gat	3461				
Lys	Tyr	Tyr	Ala	Asp	Ser	Val	Lys	Gly	Arg	Phe	Thr	Ile	Ser	Arg	Asp					
	170					175					180									
aac	agc	aag	aac	acc	ctg	tac	ctg	cag	atg	aac	agc	ctg	cgg	gcc	gag	3509				
Asn	Ser	Lys	Asn	Thr	Leu	Tyr	Leu	Gln	Met	Asn	Ser	Leu	Arg	Ala	Glu					
185					190					195				200						
gat	acc	gcc	gtg	tac	tac	tgc	gcc	aag	gat	agc	cag	ctg	cgg	agc	ctg	3557				
Asp	Thr	Ala	Val	Tyr	Tyr	Cys	Ala	Lys	Asp	Ser	Gln	Leu	Arg	Ser	Leu					
				205					210					215						
ctg	tac	ttc	gag	tgg	ctg	agc	cag	gga	tac	ttc	gat	tac	tgg	gga	cag	3605				
Leu	Tyr	Phe	Glu	Trp	Leu	Ser	Gln	Gly	Tyr	Phe	Asp	Tyr	Trp	Gly	Gln					
			220					225					230							
gga	acc	ctg	gtg	acc	gtg	agc	agc	gct	agc	acc	aag	gga	cca	agc	gtg	3653				
Gly	Thr	Leu	Val	Thr	Val	Ser	Ser	Ala	Ser	Thr	Lys	Gly	Pro	Ser	Val					
		235					240					245								
ttc	cca	ctg	gcc	cca	agc	agc	aag	agc	acc	agc	gga	gga	acc	gcc	gcc	3701				
Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys	Ser	Thr	Ser	Gly	Gly	Thr	Ala	Ala					
	250					255					260									
ctg	gga	tgc	ctg	gtg	aag	gat	tac	ttc	cca	gag	cca	gtg	acc	gtg	agc	3749				
Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	Phe	Pro	Glu	Pro	Val	Thr	Val	Ser					
265					270					275				280						
tgg	aac	agc	gga	gcc	ctg	acc	agc	gga	gtg	cac	acc	ttc	cca	gcc	gtg	3797				
Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	Gly	Val	His	Thr	Phe	Pro	Ala	Val					
				285				290						295						
ctg	cag	agc	agc	gga	ctg	tat	agc	ctg	agc	agc	gtg	gtg	acc	gtg	cca	3845				
Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val	Thr	Val	Pro					
			300					305					310							
agc	agc	agc	ctg	gga	acc	cag	acc	tac	atc	tgc	aac	gtg	aac	cac	aag	3893				
Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	Cys	Asn	Val	Asn	His	Lys					
		315					320					325								
cca	agc	aac	acc	aag	gtg	gat	aag	aag	gtg	gag	cca	aag	agc	tgc	gat	3941				
Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	Glu	Pro	Lys	Ser	Cys	Asp					
		330				335					340									
aag	acc	cac	acg	tgc	cct	cct	tgt	cca	gcc	ccc	gaa	ctg	ctg	ggc	ggg	3989				
Lys	Thr	His	Thr	Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly					
345					350					355				360						

cct agc gtg ttc ctg ttt ccc cct aag cct aaa gat aca ctg atg att	4037
Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile	
365 370 375	
agt aga acc cca gag gtc aca tgc gtg gtc gtg gac gtg tcc cac gaa	4085
Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu	
380 385 390	
gag cct gac gtg aag ttc aac tgg tac gtg gat ggc gtg gag gtg cac	4133
Glu Pro Asp Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His	
395 400 405	
aat gct aag act aaa cca cgc gaa gag cag tat aat agt aca tac cga	4181
Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg	
410 415 420	
gtc gtg tca gtc ctg aca gtg ctg cac cag gat tgg ctg aac ggc aag	4229
Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys	
425 430 435 440	
gag tat aag tgc aag gtg tct aac aag gcc ctg ccc gcc cct atc gag	4277
Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu	
445 450 455	
aaa aca att agc aag gcc aaa ggg cag cca cgg gaa ccc cag gtc tac	4325
Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr	
460 465 470	
act ctg cca ccc tca aga gat gaa ctg act aag aac cag gtc agc ctg	4373
Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu	
475 480 485	
acc tgt ctg gtg aaa ggc ttc tac ccc agc gac atc gcc gtg gag tgg	4421
Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp	
490 495 500	
gaa agt aac ggc cag cct gag aat aac tac aag act acc cct cca gtg	4469
Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val	
505 510 515 520	
ctg gat agc gac ggg tcc ttc ttc ctg tat agc aag ctg aca gtg gac	4517
Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp	
525 530 535	
aaa tcc cgc tgg cag cag gga aac gtc ttt tcc tgt tct gtg atg cat	4565
Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His	
540 545 550	
gag gcc ctg cac aat cat tac acc cag aag agt ctg tca ctg agc ccc	4613
Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro	
555 560 565	
ggc aaa tgataaaagg aacccgcgct atgacggcaa taaaaagaca gaataaaacc	4669
Gly Lys	

570

cacgggtggt	gggtcgtttg	ttcataaacc	cgggatcgat	aaggatcttc	ctagagcatg	4729
gctacgtaga	taagtagcat	ggcgggttaa	tcattaacta	caaggaaccc	ctagtgatgg	4789
agttggccac	tccctctctg	cgcgctcgct	cgctcactga	ggccgggcga	ccaaaggtcg	4849
cccgacgcc	gggctttgcc	cgggcggcct	cagtgaagca	gcgagcgcg	agccttaatt	4909
aacctaattc	actggccgtc	gttttacaac	gtcgtgactg	ggaaaaccct	ggcgttaccc	4969
aacttaatcg	ccttgacgca	catccccctt	tcgccagctg	gcgtaatagc	gaagaggccc	5029
gcaccgatcg	cccttcccaa	cagttgcgca	gcctgaatgg	cgaatgggac	gcgccctgta	5089
gcggcgcat	aagcgcggcg	ggtgtggtgg	ttacgcgcag	cgtgaccgct	acacttgcca	5149
gcgccctagc	gcccgctcct	ttcgctttct	tcccttcctt	tctcgccacg	ttcgccggct	5209
ttccccgtca	agctctaaat	cgggggctcc	ctttagggtt	ccgatttagt	gctttacggc	5269
acctcgaccc	caaaaaactt	gattaggggtg	atggttcacg	tagtgggcca	tcgccctgat	5329
agacggtttt	tcgccctttg	acgttgaggt	ccacgttctt	taatagtgga	ctcttggtcc	5389
aaactggaac	aacactcaac	cctatctcgg	tctattcttt	tgatttataa	gggattttgc	5449
cgatttcggc	ctattgggtta	aaaaatgagc	tgatttaaca	aaaatttaac	gcgaatttta	5509
acaaaatatt	aacgcttaca	atttaggtgg	cacttttcgg	ggaaatgtgc	gcggaacccc	5569
tatttgttta	tttttctaaa	tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	5629
ataaatgctt	caataatatt	gaaaaaggaa	gagtatgagt	attcaacatt	tcggtgtcgc	5689
ccttattccc	ttttttgcgg	cattttgcct	tcctgttttt	gctcaccag	aaacgctggt	5749
gaaagtaaaa	gatgctgaag	atcagttggg	tgacagagtg	ggttacatcg	aactggatct	5809
caacagcggg	aagatccttg	agagttttcg	ccccgaagaa	cgttttccaa	tgatgagcac	5869
ttttaaggtt	ctgctatgtg	gcgcggtatt	atcccgtatt	gacgccgggc	aagagcaact	5929
cggtcgccgc	atacactatt	ctcagaatga	cttggttgag	tactcaccag	tcacagaaaa	5989
gcatcttacg	gatggcatga	cagtaagaga	attatgcagt	gctgccataa	ccatgagtga	6049
taacactgcy	gccaaacttac	ttctgacaac	gatcggagga	ccgaaggagc	taaccgcttt	6109
tttgacaaac	atgggggatc	atgtaactcg	ccttgatcgt	tggaaccggg	agctgaatga	6169
agccatacca	aacgacgagc	gtgacaccac	gatgcctgta	gcaatggcaa	caacgttgcy	6229

```

caaactatta actggcgaac tacttactct agcttcccgg caacaattaa tagactggat 6289
ggaggcggat aaagttgcag gaccacttct gcgctcggcc cttccggctg gctggtttat 6349
tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt atcattgcag cactggggcc 6409
agatggtaag ccctcccgta tcgtagttat ctacacgacg gggagtcagg caactatgga 6469
tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc 6529
agaccaagtt tactcatata tacttttagat tgatttaaaa cttcattttt aatttaaaag 6589
gatctaggtg aagatccttt ttgataatct catgaccaa atcccttaac gtgagttttc 6649
gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt 6709
tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt 6769
gccggatcaa gagctacca ctctttttcc gaaggtaact ggcttcagca gagcgcagat 6829
accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga actctgtagc 6889
accgcctaca tacctcgctc tgctaatacct gttaccagtg gctgctgcca gtggcgataa 6949
gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg 7009
ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag 7069
atacctacag cgtgagctat gagaaagcgc cacgcttccc gaaggagaga aggcggacag 7129
gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa 7189
cgcttggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt 7249
gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg ccttttttacg 7309
gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc 7369
tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac 7429
cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct 7489
ccccgcgcgt tggccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc 7549
gggcagtgag cgcaacgcaa ttaatgtgag ttagctcact cattaggcac ccaggcctt 7609
acactttatg cttccggctc gtatgttggtg tggaattgtg agcggataac aatttcacac 7669
aggaaacagc tatgaccatg attacgccag atttaattaa ggccttaatt agg 7722

```

```

<210> 4
<211> 111
<212> PRT

```

<213> Trình tự nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 4

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Thr Phe Asn
20 25 30

Tyr Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro
35 40 45

Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp
50 55 60

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser
65 70 75 80

Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln His Tyr
85 90 95

Arg Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
100 105 110

<210> 5

<211> 129

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 5

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr
20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val

35						40						45					
Ala	Val	Ile	Ser	Tyr	Asp	Ala	Asn	Tyr	Lys	Tyr	Tyr	Ala	Asp	Ser	Val		
50			55			60											
Lys	Gly	Arg	Phe	Thr	Ile	Ser	Arg	Asp	Asn	Ser	Lys	Asn	Thr	Leu	Tyr		
65		70				75				80							
Leu	Gln	Met	Asn	Ser	Leu	Arg	Ala	Glu	Asp	Thr	Ala	Val	Tyr	Tyr	Cys		
85				90				95									
Ala	Lys	Asp	Ser	Gln	Leu	Arg	Ser	Leu	Leu	Tyr	Phe	Glu	Trp	Leu	Ser		
100			105			110											
Gln	Gly	Tyr	Phe	Asp	Tyr	Trp	Gly	Gln	Gly	Thr	Leu	Val	Thr	Val	Ser		
115		120				125											

Ser

```
<210> 6
<211> 107
<212> PRT
<213> Trình tự nhân tạo
```

<220>
<223> Cấu trúc tổng hợp

<400> 6

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

65	70								75								80							
Tyr	Ile	Cys	Asn	Val	Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys									
			85						90					95										
Lys	Val	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His														
			100					105																
<210>	7																							
<211>	223																							
<212>	PRT																							
<213>	Trình tự nhân tạo																							
<220>																								
<223>	Cấu trúc tổng hợp																							
<400>	7																							
Thr	Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val									
1				5				10						15										
Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr									
			20					25					30											
Pro	Glu	Val	Thr	Cys	Val	Val	Val	Asp	Val	Ser	His	Glu	Glu	Pro	Asp									
		35					40					45												
Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys									
	50					55					60													
Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser									
65				70						75					80									
Val	Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys									
				85					90					95										
Cys	Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr	Ile									
			100					105					110											
Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu	Pro									
		115					120					125												
Pro	Ser	Arg	Asp	Glu	Leu	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys	Leu									

130 135 140
 Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
 145 150 155 160
 Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
 165 170 175
 Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
 180 185 190
 Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
 195 200 205
 His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 210 215 220

<210> 8
 <211> 7773
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Plasmit mã hóa các kháng thể đơn dòng FI6 và C05

<220>
 <221> tín hiệu_polyA
 <222> (204)..(252)
 <223> tổng hợp_polyA

<220>
 <221> đặc điểm đa dạng
 <222> (259)..(927)
 <223> bổ sung- CH'2-3

<220>
 <221> đặc điểm đa dạng
 <222> (928)..(1248)
 <223> bổ sung- CH'1

<220>
 <221> đặc điểm đa dạng
 <222> (1251)..(1668)
 <223> bổ sung- C05_VH

<220>
 <221> đặc điểm đa dạng

<222> (1669)..(1719)
 <223> bổ sung- đoạn dẫn

<220>
 <221> đặc điểm đa dạng
 <222> (1729)..(1979)
 <223> bổ sung- CMV\mp2

<220>
 <221> đặc điểm đa dạng
 <222> (1798)..(2266)
 <223> Đoạn tăng cường

<220>
 <221> đặc điểm đa dạng
 <222> (2267)..(2392)
 <223> CMV\mp2

<220>
 <221> CDS
 <222> (2509)..(2841)
 <223> FI6\VL

<220>
 <221> CDS
 <222> (2842)..(3162)
 <223> CL

<220>
 <221> đặc điểm đa dạng
 <222> (3247)..(3306)
 <223> đoạn dẫn

<220>
 <221> CDS
 <222> (3307)..(3693)
 <223> FI6\VH

<220>
 <221> CDS
 <222> (3694)..(4014)
 <223> CH1

<220>
 <221> CDS
 <222> (4015)..(4683)
 <223> CH2-3

<220>
 <221> tín hiệu polyA
 <222> (4690)..(4767)
 <223> TKpA ngắn

<400> 8

ggccttaatt aggctgcgcg ctcgctcgct cactgaggcc gcccgggcaa agccccggcg	60
tcgggcgacc tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc	120
caactccatc actaggggtt ccttgtagtt aatgattaac ccgccatgct acttatctac	180
gtagccatgc tctaggaaga tctcacacaa aaaaccaaca cacagatgta atgaaaataa	240
agatatttta ttttatcact tcccggggct caggctcagg gacttctggg tgtagtggtt	300
gtgcagggcc tcgtgcatca cgctgcagct gaacacgttg ccctgctgcc accggctcct	360
gtccacggtc agcttgctat acaggaagaa tgagccgtcg ctgtccagca cagggggggg	420
ggtctttagt ttgttctcgg gctggccgtt gctctcccat tccacggcga tctcgctggg	480
gtagaagccc ttgaccaggc aggtcagggg cacctggttc ttggtcatct cttcccggct	540
ggggggcagt gtgtagacct gaggtcgcgc gggctggccc ttggccttgc tgatggtttt	600
ctcgatgggg gcaggcaggg ccttgttgga caccttgcac ttgtactctt tgccgttcag	660
ccagtcctgg tgcagcacgg tcagcacgga caccaccgg taggtgctgt tgtactgttc	720
ctctctgggc ttggtcttgg cgttggtgcac ttccacgccg tccacgtacc aattgaactt	780
cacttcaggg tctcgtggg acacgtccac caccacgcag gtcacttcgg gggtcgggt	840
gatcatcagg gtgtccttgg gctttggggg gaacaggaac acgctggggc ctcccagcag	900
ttcaggggca gggcaggggg gacacgtgtg ggtcttgtcg cagctcttag gttccaccgc	960
cttgtccacc ttggtgttgc tgggcttgtg gttcacgttg cagatgtagg tctgggtgcc	1020
caggctgctg ctgggcacgg tgaccacgct gtcaggcta tacaggccgc tgctctgcag	1080
cacggctgga aagggtgtgca cgccgctggg cagggcgcca gagttccagg acacggtcac	1140
gggctcgggg aagtagtcct tgaccaggca gcccagggcg gctgttccgc cagagggtgt	1200
cttgctgcta ggggccagag ggaacacgct tggtcccttg gtgctggcgc tcgagacggt	1260
caccaggggt ccctgtcccc acacatcgaa ggcattctcc accagatcgg cccgctccca	1320
tccggcgctc accacctgct gcatggacat gtgcttggcg cagtagtaca ctccggtatc	1380
ctccaccgc aggttgggtca tctgcaggta cagggtctcc ttgctgttat cccggctgat	1440
ggtgaaccgt ccctccacgc tatcggcgta atcaatgtct cctcctccgg cgttgatgat	1500
gctcagccac tccagtcctt ttccctggggc ctgccgcacc cagctcacgg cgtagtagct	1560
cagggtgctc tctccgaagc tgcttccgct tcccacgcag ctcagccgca ggctctctcc	1620

tggctgcacc agtcctcctc cgctctcctg cagctgcacc tgtgaattcg tcaccagggc	1680
caggctcagg gcgatcagca gcagcagctg catgcgcatg gtggcggcgc gatctgacgg	1740
ttcactaaac gagctctgct tatataggcc tcccaccgta cacgccacct cgacataacct	1800
agttattaat agtaatcaat tacgggggtca ttagttcata gcccatatat ggagttccgc	1860
gttacataac ttacggtaaa tggcccgctt ggctgaccgc ccaacgaccc ccgcccattg	1920
acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa	1980
tggggtggagt atttacggta aactgcccac ttggcagtac atcaagtgtg tcatatgccg	2040
agtacgcccc ctattgacgt caatgacggt aaatggcccg cctggcatta tgcccagtac	2100
atgaccttat gggactttcc tacttggcag tacatctacg tattagtcac cgctattacc	2160
atgggtgatgc ggttttggca gtacatcaat gggcgtggat agcggtttga ctcacgggga	2220
tttccaagtc tccaccccat tgacgtcaat gggagtttgt tttggcacca aaatcaacgg	2280
gactttccaa aatgtcgtaa caactccgcc ccattgacgc aaatgggcgg taggcgtgta	2340
cggtgggagg tctatataag cagagctggt ttagtgaacc gtcagatccg ctgggcactt	2400
tgcactggaa cttacaacac ccgagcaagg acgcgactct gccgccccac catgcgcatg	2460
cagctgctgc tgctgatcgc cctgagcctg gccctggtga ccaacagc gat atc gtc	2517
	Asp Ile Val
	1
atg acc cag agc cca gat agc ctg gcc gtg agc ctg gga gag cgg gcc	2565
Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly Glu Arg Ala	
5 10 15	
acc atc aac tgc aag agc agc cag agc gtg acc ttc aac tac aag aac	2613
Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Thr Phe Asn Tyr Lys Asn	
20 25 30 35	
tac ctg gcc tgg tac cag cag aag cca gga cag cca cca aag ctg ctg	2661
Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu	
40 45 50	
atc tac tgg gcc agc acc cgg gag agc gga gtg cca gat cgg ttc agc	2709
Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg Phe Ser	
55 60 65	
gga agc gga agc gga acc gat ttc acc ctg acc atc agc agc ctg cag	2757
Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln	
70 75 80	
gcc gag gat gtg gcc gtg tac tac tgc cag cag cac tac cgg acc cca	2805
Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln His Tyr Arg Thr Pro	

85	90	95	
cca acc ttc gga cag gga acc aag gtg gag atc aag cgt acg gtg gcc			2853
Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala			
100	105	110	115
gcc cca agc gtg ttc atc ttc cca cca agc gat gag cag ctg aag agc			2901
Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser			
	120	125	130
gga acc gcc agc gtg gtg tgc ctg ctg aac aac ttc tac cca cgg gag			2949
Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu			
	135	140	145
gcc aag gtg cag tgg aag gtg gat aac gcc ctg cag agc gga aac agc			2997
Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser			
	150	155	160
cag gag agc gtg acc gag cag gat agc aag gat agc acc tac agc ctg			3045
Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu			
	165	170	175
agc agc acc ctg acc ctg agc aag gcc gat tac gag aag cac aag gtg			3093
Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val			
180	185	190	195
tac gcc tgc gag gtg acc cac cag gga ctg agc agc cca gtg acc aag			3141
Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys			
	200	205	210
agc ttc aac cgc gga gag tgc cggaagcggc gggccccagt gaagcagacc			3192
Ser Phe Asn Arg Gly Glu Cys			
	215		
ctgaacttcg atctgctgaa gctggccgga gatgtggaga gcaaccagga accaatgtac			3252
agaatgcagc tgctgagctg catcgccctg agcctggccc tggtagaccaa cagc cag			3309
			Gln
gtg caa cta gtg gag agc gga gga gga gtg gtg cag cca gga cgg agc			3357
Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg Ser			
220	225	230	235
ctg cgg ctg agc tgc gcc gcc agc gga ttc acc ttc agc acc tac gcc			3405
Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr Ala			
	240	245	250
atg cac tgg gtg cgg cag gcc cca gga aag gga ctg gag tgg gtg gcc			3453
Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ala			
	255	260	265
gtg atc agc tac gat gcc aac tac aag tac tac gcc gat agc gtg aag			3501
Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser Val Lys			

270	275	280	
gga cgg ttc acc atc agc cgg gat aac agc aag aac acc ctg tac ctg			3549
Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu			
285	290	295	
cag atg aac agc ctg cgg gcc gag gat acc gcc gtg tac tac tgc gcc			3597
Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala			
300	305	310	315
aag gat agc cag ctg cgg agc ctg ctg tac ttc gag tgg ctg agc cag			3645
Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu Ser Gln			
	320	325	330
gga tac ttc gat tac tgg gga cag gga acc ctg gtg acc gtg agc agc			3693
Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser			
	335	340	345
gct agc acc aag gga cca agc gtg ttc cca ctg gcc cca agc agc aag			3741
Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys			
	350	355	360
agc acc agc gga gga acc gcc gcc ctg gga tgc ctg gtg aag gat tac			3789
Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr			
	365	370	375
ttc cca gag cca gtg acc gtg agc tgg aac agc gga gcc ctg acc agc			3837
Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser			
380	385	390	395
gga gtg cac acc ttc cca gcc gtg ctg cag agc agc gga ctg tat agc			3885
Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser			
	400	405	410
ctg agc agc gtg gtg acc gtg cca agc agc agc ctg gga acc cag acc			3933
Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr			
	415	420	425
tac atc tgc aac gtg aac cac aag cca agc aac acc aag gtg gat aag			3981
Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys			
	430	435	440
aag gtg gag cca aag agc tgc gat aag acc cac acg tgc cct cca tgt			4029
Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys			
	445	450	455
cca gcc ccc gaa ctg ctg ggc ggg cct agc gtg ttc ctg ttt ccc cct			4077
Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro			
460	465	470	475
aag cct aaa gat aca ctg atg att agt aga acc cca gag gtc aca tgc			4125
Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys			
	480	485	490

gtg gtc gtg gac gtg tcc cac gaa gag cct gac gtg aag ttc aac tgg	4173
Val Val Val Asp Val Ser His Glu Glu Pro Asp Val Lys Phe Asn Trp	
495 500 505	
tac gtg gat ggc gtg gag gtg cac aat gct aag act aaa cca cgc gaa	4221
Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu	
510 515 520	
gag cag tat aat agt aca tac cga gtc gtg tca gtc ctg aca gtg ctg	4269
Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu	
525 530 535	
cac cag gat tgg ctg aac ggc aag gag tat aag tgc aag gtg tct aac	4317
His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn	
540 545 550 555	
aag gcc ctg ccc gcc cct atc gag aaa aca att agc aag gcc aaa ggg	4365
Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly	
560 565 570	
cag cca cgg gaa ccc cag gtc tac act ctg cca ccc tca aga gat gaa	4413
Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu	
575 580 585	
ctg act aag aac cag gtc agc ctg acc tgt ctg gtg aaa ggc ttc tac	4461
Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr	
590 595 600	
ccc agc gac atc gcc gtg gag tgg gaa agt aac ggc cag cct gag aat	4509
Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn	
605 610 615	
aac tac aag act acc cct cca gtg ctg gat agc gac ggg tcc ttc ttc	4557
Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe	
620 625 630 635	
ctg tat agc aag ctg aca gtg gac aaa tcc cgc tgg cag cag gga aac	4605
Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn	
640 645 650	
gtc ttt tcc tgt tct gtg atg cat gag gcc ctg cac aat cat tac acc	4653
Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr	
655 660 665	
cag aag agt ctg tca ctg agc ccc ggc aaa tgataaaagg aacccgcgct	4703
Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys	
670 675	
atgacggcaa taaaaagaca gaataaaacc cacgggtgtt gggtcgtttg ttcataaacc	4763
cgggatcgat aaggatcttc ctagagcatg gctacgtaga taagtagcat ggcggggttaa	4823
tcattaacta caaggaaccc ctagtgatgg agttggccac tccctctctg cgcgctcgct	4883

cgctcactga	ggccggggcga	ccaaagggtcg	cccgcgccc	gggctttgcc	cgggcggcct	4943
cagtgcgcga	gcgagcgcg	agccttaatt	aacctaattc	actggccgtc	gttttacaac	5003
gtcgtgactg	ggaaaaccct	ggcgttaccc	aacttaatcg	ccttgcagca	catccccctt	5063
tcgccagctg	gcgtaatagc	gaagaggccc	gcaccgatcg	cccttcccaa	cagttgcgca	5123
gcctgaatgg	cgaatgggac	gcgccttgta	gcggcgcatc	aagcgcggcg	ggtgtggtgg	5183
ttacgcgcag	cgtgaccgct	acacttgcca	gcgccttagc	gcccgcctct	ttcgctttct	5243
tcccttcctt	tctcgccaag	ttcgccggct	ttccccgtca	agctctaaat	cgggggctcc	5303
ctttagggtt	ccgatttagt	gctttacggc	acctcgaccc	caaaaaactt	gatttagggtg	5363
atggttcacg	tagtgggcca	tcgccctgat	agacggtttt	tcgccctttg	acgttggagt	5423
ccacgttctt	taatagtgga	ctcttgttcc	aaactggaac	aacactcaac	cctatctcgg	5483
tctattcttt	tgatttataa	gggattttgc	cgatttcggc	ctattggtta	aaaaatgagc	5543
tgatttaaca	aaaattttaac	gcgaatttta	acaaaatatt	aacgcttaca	atttaggtgg	5603
cacttttcgg	ggaaatgtgc	gcggaacccc	tatttgttta	tttttctaaa	tacattcaaa	5663
tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	caataatatt	gaaaaaggaa	5723
gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	ttttttgcgg	cattttgcct	5783
tcctgttttt	gctcaccag	aaacgctggg	gaaagtaaaa	gatgctgaag	atcagttggg	5843
tgcacgagtg	ggttacatcg	aactggatct	caacagcggg	aagatccttg	agagttttcg	5903
ccccgaagaa	cgtttttcaa	tgatgagcac	ttttaaggtt	ctgctatgtg	gcgcggtatt	5963
atcccgtatt	gacgcggggc	aagagcaact	cggtcgccgc	atacactatt	ctcagaatga	6023
cttggttgag	tactcaccag	tcacagaaaa	gcattcttac	gatggcatga	cagtaagaga	6083
attatgcagt	gctgccataa	ccatgagtga	taacactgcg	gccaacttac	ttctgacaac	6143
gatcggagga	ccgaaggagc	taaccgcttt	tttgcacaa	atgggggatc	atgtaactcg	6203
ccttgatcgt	tggaaccggg	agctgaatga	agccatacca	aacgacgagc	gtgacaccac	6263
gatgcctgta	gcaatggcaa	caacgttgcg	caaactatta	actggcgaac	tacttactct	6323
agcttcccgg	caacaattaa	tagactggat	ggaggcggat	aaagttgcag	gaccacttct	6383
gcgctcggcc	cttcgggctg	gctggtttat	tgctgataaa	tctggagccg	gtgagcgtgg	6443
gtctcgcggg	atcattgcag	cactggggcc	agatggtaag	ccctcccgta	tcgtagttat	6503

ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg 6563
 tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tacttttagat 6623
 tgatttaaaa cttcattttt aattttaaag gatctagggtg aagatccttt ttgataatct 6683
 catgacaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 6743
 gatcaaagga tcttcttgag atcctttttt totgcgcgta atctgctgct tgcaaacaaa 6803
 aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa gagctaccaa ctctttttcc 6863
 gaaggtaact ggcttcagca gagcgcagat accaaatact gttcttctag tgtagccgta 6923
 gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaatact 6983
 gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttg actcaagacg 7043
 atagttaccg gataaggcgc agcggtcggg ctgaacgggg gggttcgtgca cacagcccag 7103
 cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc 7163
 cacgcttccc gaaggagaaa aggcggacag gtatccggta agcggcaggg tcggaacagg 7223
 agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc ctgtcggggtt 7283
 tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg 7343
 gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc cttttgctca 7403
 catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg 7463
 agctgatacc gctcgcgcga gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc 7523
 ggaagagcgc ccaatacgca aaccgcctct ccccgcgctt tggccgattc attaatgcag 7583
 ctggcacgac aggtttcccg actggaaaagc gggcagtgag cgcaacgcaa ttaatgtgag 7643
 ttagctcact cattaggcac ccaggtttt acactttatg cttccggctc gtatgttggtg 7703
 tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg attacgccag 7763
 atttaattaa 7773

<210> 9
 <211> 111
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 9

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Thr Phe Asn
20 25 30

Tyr Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro
35 40 45

Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp
50 55 60

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser
65 70 75 80

Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln His Tyr
85 90 95

Arg Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
100 105 110

<210> 10

<211> 107

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 10

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
1 5 10 15

Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe
20 25 30

Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln
35 40 45

Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
50 55 60

Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
65 70 75 80

Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
85 90 95

Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
100 105

<210> 11
<211> 129
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc tổng hợp

<400> 11

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr
20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu Ser
100 105 110

Gln Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
115 120 125

Ser

<210> 12
 <211> 107
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 12

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
 1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
 20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
 35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
 50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
 65 70 75 80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
 85 90 95

Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His
 100 105

<210> 13
 <211> 223
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 13

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val

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

<210> 14
 <211> 7728
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Plasmit mã hóa các kháng thể đơn dòng FI6 và 1A6 tổng hợp

<220>
 <221> tín hiệu_polyA
 <222> (191)..(239)
 <223> tổng hợp_polyA

<220>
 <221> đặc điểm_đa dạng
 <222> (246)..(914)
 <223> bổ sung- CH'2-3

<220>
 <221> đặc điểm_đa dạng
 <222> (915)..(1235)
 <223> bổ sung- CH'1

<220>
 <221> đặc điểm_đa dạng
 <222> (1236)..(1598)
 <223> bổ sung- 1A6_VH

<220>
 <221> đặc điểm_đa dạng
 <222> (1599)..(1655)
 <223> bổ sung- đoạn dẫn

<220>
 <221> đặc điểm_đa dạng
 <222> (1665)..(1733)
 <223> bổ sung- CMV_mp2

<220>
 <221> đặc điểm_đa dạng
 <222> (1732)..(2202)
 <223> Đoạn tăng cường

<220>
 <221> đặc điểm_đa dạng
 <222> (2203)..(2328)
 <223> CMV_mp1

<220>
 <221> đặc điểm_đa dạng
 <222> (2388)..(2444)
 <223> đoạn dẫn

<220>
 <221> đặc điểm đa dạng
 <222> (2445) .. (2789)
 <223> KGL

<220>
 <221> CDS
 <222> (2784) .. (3104)
 <223> CL

<220>
 <221> đặc điểm đa dạng
 <222> (3189) .. (3248)
 <223> đoạn dẫn

<220>
 <221> CDS
 <222> (3249) .. (3635)
 <223> FI6\ VH

<220>
 <221> CDS
 <222> (3636) .. (3956)
 <223> CH1

<220>
 <221> CDS
 <222> (3957) .. (4625)
 <223> CH2-3

<220>
 <221> tín hiệu polyA
 <222> (4632) .. (4709)
 <223> TKpA ngắn

<400> 14
 ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgctcg ggcgaccttt 60
 ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120
 aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct 180
 aggaagatct cacacaaaa accaacacac agatgtaatg aaaataaaga tattttatatt 240
 tatcacttcc cggggctcag gctcagggaac ttctgggtgt agtgggtgtg cagggcctcg 300
 tgcattcacgc tgcagctgaa cacgttgccc tgctgccacc ggctcttgtc cacggtcagc 360
 ttgctataca ggaagaatga gccgtcgctg tccagcacag ggggggtggt cttgtagttg 420
 ttctcgggct ggccgttgct ctccattcc acggcgatct cgctggggta gaagcccttg 480
 accaggcagg tcagggaacac ctggttcttg gtcattctctt cccggctggg gggcagtggtg 540

tagacctgag gctcgcgggg ctggcccttg gccttgctga tggttttctc gatgggggca	600
ggcagggcct tgttggaacac cttgcacttg tactctttgc cgttcagcca gtcctggtgc	660
agcacggtca gcacggacac cacccggtag gtgctgttgt actgttcctc tctgggcttg	720
gtcttggcgt tgtgcacttc cacgccgtcc acgtaccaat tgaacttcac ttcagggtec	780
tcgtgggaca cgtccaccac cacgcaggtc acttcggggg tccggctgat catcagggtg	840
tccttgggct ttggggggaa caggaacacg ctggggcctc ccagcagttc aggggcaggg	900
caggggggac acgtgtgggt cttgtcgcag ctcttaggtt ccaccgctt gtccaccttg	960
gtgttgctgg gcttgtgggt cacgttgacg atgtaggtct gggtgcccag gctgctgctg	1020
ggcacggtga ccacgtgct caggctatac aggccgtgc tctgcagcac ggctggaaag	1080
gtgtgcacgc cgctgggtcag ggcgccagag ttccaggaca cggtcacggg ctcggggaag	1140
tagtccttga ccaggcagcc cagggcggct gttccgccag aggtgctctt gctgctaggg	1200
gccagaggga acacgcttgg tcccttggtg ctggcgctcg agacggtcac cagggttccc	1260
tgtccccagt aatccattcc tccgtggcg attccgctcc gatccttggc gcagtagtac	1320
acggcggtat cctcggcccg caggctgttc atctgcaggt acagggtggt cttgctgttg	1380
gcccggtga tgggtgaaccg tcccttcacg ctatcggcgt agtacttggt gtttccatcg	1440
tagctgatca cggccacca ctccagtccc tttcctgggg cctgccgcac ccagtgcatt	1500
ccgtaatcgc tgaaggtgaa tccgtggcg gcgcagctca gccgcaggct ccgtcctggc	1560
tgcaccactc ctctccgct ctctgcagc tgcacctgtg aattcgtcac cagggccagg	1620
ctcagggcga tcagcagcag cagctgcatg cgcattggtg cggcgcgatc tgacggttca	1680
ctaaacgagc tctgcttata taggcctccc accgtacacg ccacctcgac atacctagtt	1740
attaatagta atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgtta	1800
cataacttac ggtaaatggc ccgcctggct gaccgcccaa cgacccccgc ccattgacgt	1860
caataatgac gtatgttccc atagtaacgc caataggac tttccattga cgtcaatggg	1920
tggagtattht acggtaaact gccacttg cagtacatca agtgtatcat atgccaagta	1980
cgccccctat tgacgtcaat gacggtaaht ggccgcctg gcattatgcc cagtacatga	2040
ccttatggga ctttcctact tggcagtaca tctacgtatt agtcatcgct attaccatgg	2100
tgatgcgggtt ttggcagtag atcaatgggc gtggatagcg gtttgactca cggggatttc	2160

caagtctcca cccattgac gtcaatggga gtttgttttg gcacaaaat caacgggact	2220
ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat gggcggtagg cgtgtacggt	2280
gggaggtcta tataagcaga gctggtttag tgaaccgtca gatccgctgg gcactttgca	2340
ctggaactta caacacccga gcaaggacgc gactctgccg cccaccatg cgcattgcagc	2400
tgctgctgct gatcgccctg agcctggccc tggtagcaaa cagcgatatc gtcattgacct	2460
agagcccaga tagcctggcc gtgagcctgg gagagcgggc caccatcaac tgcaagagca	2520
gccagagcgt gctgtacagc agcaacaaca agaactacct ggccctggtac cagcagaagc	2580
caggacagcc accaaagctg ctgatctact gggccagcac ccgggagagc ggagtgccag	2640
atcggttcag cggaagcggg agcgggaaccg atttcaccct gaccatcagc agcctgcagg	2700
ccgaggatgt ggccgtgtac tactgccagc agtactacag caccctactg accttcggac	2760
agggaaccaa ggtggagatc aag cgt acg gtg gcc gcc cca agc gtg ttc atc	2813
Arg Thr Val Ala Ala Pro Ser Val Phe Ile	
1 5 10	
ttc cca cca agc gat gag cag ctg aag agc gga acc gcc agc gtg gtg	2861
Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val	
15 20 25	
tgc ctg ctg aac aac ttc tac cca cgg gag gcc aag gtg cag tgg aag	2909
Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys	
30 35 40	
gtg gat aac gcc ctg cag agc gga aac agc cag gag agc gtg acc gag	2957
Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu	
45 50 55	
cag gat agc aag gat agc acc tac agc ctg agc agc acc ctg acc ctg	3005
Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu	
60 65 70	
agc aag gcc gat tac gag aag cac aag gtg tac gcc tgc gag gtg acc	3053
Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr	
75 80 85 90	
cac cag gga ctg agc agc cca gtg acc aag agc ttc aac cgc gga gag	3101
His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu	
95 100 105	
tgc cggaagcggc gggccccagt gaagcagacc ctgaacttcg atctgctgaa	3154
Cys	
gctggccgga gatgtggaga gcaacccagg accaatgtac agaattgcagc tgctgagctg	3214

catcgccctg agcctggccc tggtagaccaa cagc cag gtg caa cta gtg gag agc	3269
Gln Val Gln Leu Val Glu Ser	
110	
gga gga gga gtg gtg cag cca gga cgg agc ctg cgg ctg agc tgc gcc	3317
Gly Gly Gly Val Val Gln Pro Gly Arg Ser Leu Arg Leu Ser Cys Ala	
115 120 125 130	
gcc agc gga ttc acc ttc agc acc tac gcc atg cac tgg gtg cgg cag	3365
Ala Ser Gly Phe Thr Phe Ser Thr Tyr Ala Met His Trp Val Arg Gln	
135 140 145	
gcc cca gga aag gga ctg gag tgg gtg gcc gtg atc agc tac gat gcc	3413
Ala Pro Gly Lys Gly Leu Glu Trp Val Ala Val Ile Ser Tyr Asp Ala	
150 155 160	
aac tac aag tac tac gcc gat agc gtg aag gga cgg ttc acc atc agc	3461
Asn Tyr Lys Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser	
165 170 175	
cgg gat aac agc aag aac acc ctg tac ctg cag atg aac agc ctg cgg	3509
Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg	
180 185 190	
gcc gag gat acc gcc gtg tac tac tgc gcc aag gat agc cag ctg cgg	3557
Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Lys Asp Ser Gln Leu Arg	
195 200 205 210	
agc ctg ctg tac ttc gag tgg ctg agc cag gga tac ttc gat tac tgg	3605
Ser Leu Leu Tyr Phe Glu Trp Leu Ser Gln Gly Tyr Phe Asp Tyr Trp	
215 220 225	
gga cag gga acc ctg gtg acc gtg agc agc gct agc acc aag gga cca	3653
Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro	
230 235 240	
agc gtg ttc cca ctg gcc cca agc agc aag agc acc agc gga gga acc	3701
Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr	
245 250 255	
gcc gcc ctg gga tgc ctg gtg aag gat tac ttc cca gag cca gtg acc	3749
Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr	
260 265 270	
gtg agc tgg aac agc gga gcc ctg acc agc gga gtg cac acc ttc cca	3797
Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro	
275 280 285 290	
gcc gtg ctg cag agc agc gga ctg tat agc ctg agc agc gtg gtg acc	3845
Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr	
295 300 305	
gtg cca agc agc agc ctg gga acc cag acc tac atc tgc aac gtg aac	3893

Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	Cys	Asn	Val	Asn		
			310					315					320				
cac	aag	cca	agc	aac	acc	aag	gtg	gat	aag	aag	gtg	gag	cca	aag	agc		3941
His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	Glu	Pro	Lys	Ser		
		325					330					335					
tgc	gat	aag	acc	cac	acg	tgc	cct	cct	tgt	cca	gcc	ccc	gaa	ctg	ctg		3989
Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu		
	340					345					350						
ggc	ggg	cct	agc	gtg	ttc	ctg	ttt	ccc	cct	aag	cct	aaa	gat	aca	ctg		4037
Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu		
355					360					365					370		
atg	att	agt	aga	acc	cca	gag	gtc	aca	tgc	gtg	gtc	gtg	gac	gtg	tcc		4085
Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys	Val	Val	Val	Asp	Val	Ser		
				375					380					385			
cac	gaa	gag	cct	gac	gtg	aag	ttc	aac	tgg	tac	gtg	gat	ggc	gtg	gag		4133
His	Glu	Glu	Pro	Asp	Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu		
			390					395					400				
gtg	cac	aat	gct	aag	act	aaa	cca	cgc	gaa	gag	cag	tat	aat	agt	aca		4181
Val	His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr		
		405					410					415					
tac	cga	gtc	gtg	tca	gtc	ctg	aca	gtg	ctg	cac	cag	gat	tgg	ctg	aac		4229
Tyr	Arg	Val	Val	Ser	Val	Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn		
	420					425					430						
ggc	aag	gag	tat	aag	tgc	aag	gtg	tct	aac	aag	gcc	ctg	ccc	gcc	cct		4277
Gly	Lys	Glu	Tyr	Lys	Cys	Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro		
435					440					445					450		
atc	gag	aaa	aca	att	agc	aag	gcc	aaa	ggg	cag	cca	cgg	gaa	ccc	cag		4325
Ile	Glu	Lys	Thr	Ile	Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln		
				455					460					465			
gtc	tac	act	ctg	cca	ccc	tca	aga	gat	gaa	ctg	act	aag	aac	cag	gtc		4373
Val	Tyr	Thr	Leu	Pro	Pro	Ser	Arg	Asp	Glu	Leu	Thr	Lys	Asn	Gln	Val		
			470					475					480				
agc	ctg	acc	tgt	ctg	gtg	aaa	ggc	ttc	tac	ccc	agc	gac	atc	gcc	gtg		4421
Ser	Leu	Thr	Cys	Leu	Val	Lys	Gly	Phe	Tyr	Pro	Ser	Asp	Ile	Ala	Val		
		485					490					495					
gag	tgg	gaa	agt	aac	ggc	cag	cct	gag	aat	aac	tac	aag	act	acc	cct		4469
Glu	Trp	Glu	Ser	Asn	Gly	Gln	Pro	Glu	Asn	Asn	Tyr	Lys	Thr	Thr	Pro		
	500					505					510						
cca	gtg	ctg	gat	agc	gac	ggg	tcc	ttc	ttc	ctg	tat	agc	aag	ctg	aca		4517
Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	Phe	Phe	Leu	Tyr	Ser	Lys	Leu	Thr		
515					520					525					530		

gtg gac aaa tcc cgc tgg cag cag gga aac gtc ttt tcc tgt tct gtg	4565
Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val	
535 540 545	
atg cat gag gcc ctg cac aat cat tac acc cag aag agt ctg tca ctg	4613
Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu	
550 555 560	
agc ccc ggc aaa tgataaaagg aacccgcgct atgacggcaa taaaaagaca	4665
Ser Pro Gly Lys	
565	
gaataaaacc cacgggtggtt gggtcgtttg ttcataaacc cgggatcgat aaggatcttc	4725
ctagagcatg gctacgtaga taagtagcat ggcggggttaa tcattaacta caaggaaccc	4785
ctagtgatgg agttggccac tccctctctg cgcgctcgct cgctcactga ggccgggcca	4845
ccaaaggctg cccgacgccc gggctttgcc cgggcggcct cagtgagcga gcgagcgcg	4905
agccttaatt aacctaattc actggccgctc gttttacaac gtcgtgactg ggaaaaccct	4965
ggcgttaccc aacttaatcg ccttgacgca catccccctt tcgccagctg gcgtaatagc	5025
gaagaggccc gcaccgatcg cccttcccaa cagttgcgca gcctgaatgg cgaatgggac	5085
gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct	5145
acacttgcca gcgccctagc gcccgctcct ttcgctttct tcccttcctt tctcgccacg	5205
ttcgccggct ttccccgtca agctctaaat cgggggctcc ctttaggggt ccgatttagt	5265
gctttacggc acctcgaccc caaaaaactt gattaggggtg atggttcacg tagtgggcca	5325
tcgccctgat agacggtttt tcgccctttg acgttgaggt ccacgttctt taatagtgga	5385
ctcttggtcc aaactggaac aacactcaac cctatctcgg tctattcttt tgatttataa	5445
gggattttgc cgatttcggc ctattggtta aaaaatgagc tgatttaaca aaaatttaac	5505
gcgaatttta acaaaatatt aacgcttaca atttaggtgg cacttttcgg ggaaatgtgc	5565
gcggaacccc tatttggtta tttttctaaa tacattcaaa tatgtatccg ctcatgagac	5625
aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt	5685
tccgtgtcgc ccttattccc ttttttgccg cattttgcct tcctgttttt gctcaccag	5745
aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg gggtacatcg	5805
aactggatct caacagcggg aagatccttg agagttttcg ccccgaagaa cgttttccaa	5865
tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atcccgatt gacgccgggc	5925

aagagcaact	cggtcgcgcg	atacactatt	ctcagaatga	cttggttgag	tactcaccag	5985
tcacagaaaa	gcaccttacg	gatggcatga	cagtaagaga	attatgcagt	gctgccataa	6045
ccatgagtga	taacactgcg	gccaaacttac	ttctgacaac	gatcggagga	ccgaaggagc	6105
taaccgcttt	tttgacacaac	atgggggatc	atgtaactcg	ccttgatcgt	tgggaaccgg	6165
agctgaatga	agccatacca	aacgacgagc	gtgacaccac	gatgcctgta	gcaatggcaa	6225
caacgttgcg	caaaactatta	actggcgaac	tacttactct	agcttcccgg	caacaattaa	6285
tagactggat	ggaggcggat	aaagttgcag	gaccacttct	gcgctcggcc	cttcgggctg	6345
gctggtttat	tgctgataaa	tctggagccg	gtgagcgtgg	gtctcgcggg	atcattgcag	6405
cactggggcc	agatggtaag	ccctcccgtg	tcgtagtatt	ctacacgacg	gggagtcagg	6465
caactatgga	tgaacgaaat	agacagatcg	ctgagatagg	tgccctactg	attaagcatt	6525
ggtaactgtc	agaccaagtt	tactcatata	tacttttagat	tgatttaaaa	cttcattttt	6585
aatttaaaag	gatctaggtg	aagatccttt	ttgataatct	catgaccaa	atcccttaac	6645
gtgagttttc	gttcactga	gcgtcagacc	ccgtagaaaa	gatcaaagga	tcttcttgag	6705
atcctttttt	tctgcgcgta	atctgctgct	tgcaaacaaa	aaaaccaccg	ctaccagcgg	6765
tggtttgttt	gccggatcaa	gagctaccaa	ctctttttcc	gaaggtaact	ggcttcagca	6825
gagcgcagat	accaaatact	gttcttctag	tgtagccgta	gttaggccac	cacttcaaga	6885
actctgtagc	accgcctaca	tacctcgctc	tgctaatcct	gttaccagtg	gctgctgcca	6945
gtggcgataa	gtcgtgtcct	accgggttgg	actcaagacg	atagttaccg	gataaggcgc	7005
agcggtcggg	ctgaacgggg	ggttcgtgca	cacagcccag	cttgagagca	acgacctaca	7065
ccgaactgag	atacctacag	cgtgagctat	gagaaagcgc	cacgcttccc	gaagggagaa	7125
aggcggacag	gtatccggta	agcggcaggg	tcggaacagg	agagcgcacg	agggagcttc	7185
cagggggaaa	cgcttggtat	ctttatagtc	ctgtcggggt	tcgccacctc	tgacttgagc	7245
gtcgattttt	gtgatgctcg	tcaggggggc	ggagcctatg	gaaaaacgcc	agcaacgcgg	7305
ccttttttacg	gttcctggcc	ttttgctggc	cttttgctca	catgttcttt	cctgcgttat	7365
cccctgattc	tgtggataac	cgtattaccg	cctttgagtg	agctgatacc	gctcgcgcga	7425
gccgaacgac	cgagcgcagc	gagtcagtga	gcgaggaagc	ggaagagcgc	ccaatacgca	7485
aaccgcctct	ccccgcgcgt	tggccgattc	attaatgcag	ctggcacgac	aggtttcccg	7545

actggaaagc gggcagtgag cgcaacgcaa ttaatgtgag ttagctcact cattagggcac 7605
cccaggcttt acactttatg cttccggctc gtatgttggtg tggaattgtg agcggataac 7665
aatttcacac aggaaacagc tatgaccatg attacgccag atttaattaa ggccttaatt 7725
agg 7728

<210> 15
<211> 107
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc tổng hợp

<400> 15

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
1 5 10 15

Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe
20 25 30

Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln
35 40 45

Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
50 55 60

Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
65 70 75 80

Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
85 90 95

Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
100 105

<210> 16
<211> 129
<212> PRT
<213> Trình tự nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 16

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr
20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu Ser
100 105 110

Gln Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
115 120 125

Ser

<210> 17

<211> 107

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 17

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
65 70 75 80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
85 90 95

Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His
100 105

<210> 18
<211> 223
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc tổng hợp

<400> 18

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
1 5 10 15

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
20 25 30

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Glu Pro Asp
35 40 45

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
50 55 60

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

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
85 90 95

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
100 105 110

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
115 120 125

Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
130 135 140

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
145 150 155 160

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
165 170 175

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
180 185 190

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
195 200 205

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
210 215 220

<210> 19

<211> 7746

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Plasmit mang kháng thể đơn dòng FI6 và CR8033

<220>

<221> tín hiệu_polyA

<222> (173)..(221)

<223> tổng hợp_polyA

<220>

<221> đặc điểm đa dạng

<222> (228)..(896)

<223> bổ sung- CH'2-3

<220>
 <221> đặc điểm đa dạng
 <222> (897)..(1217)
 <223> bổ sung- CH'1

<220>
 <221> đặc điểm đa dạng
 <222> (1218)..(1604)
 <223> bổ sung- CR8033\VH

<220>
 <221> đặc điểm đa dạng
 <222> (1605)..(1655)
 <223> bổ sung- đoạn dẫn

<220>
 <221> đặc điểm đa dạng
 <222> (1665)..(1733)
 <223> bổ sung- CMV\mp2

<220>
 <221> đặc điểm đa dạng
 <222> (1734)..(2202)
 <223> Đoạn tăng cường

<220>
 <221> đặc điểm đa dạng
 <222> (2203)..(2328)
 <223> CMV\mp1

<220>
 <221> đặc điểm đa dạng
 <222> (2445)..(2789)
 <223> KGL

<220>
 <221> CDS
 <222> (2784)..(3104)
 <223> CL

<220>
 <221> đặc điểm đa dạng
 <222> (3189)..(3248)
 <223> đoạn dẫn

<220>
 <221> CDS
 <222> (3249)..(3635)
 <223> FI6\VH

<220>
 <221> CDS
 <222> (3636)..(3956)

<223> CH1

<220>

<221> CDS

<222> (3957)..(4625)

<223> CH2-3

<220>

<221> đặc điểm đa dạng

<222> (3968)..(3968)

<223> A -> T

<220>

<221> tín hiệu polyA

<222> (4632)..(4709)

<223> TKpA ngắn

<400> 19

actgaggccg cccggggcaaa gcccgggctg cgggagacct ttgggtcgccc ggcctcagtg	60
agcgagcgag cgcgagagaga gggagtggcc aactccatca ctagggggttc cttgtagtta	120
atgattaacc cgccatgcta cttatctacg tagccatgct ctagggaagat ctcacacaaa	180
aaaccaacac acagatgtaa tgaaaataaaa gatattttat tttatcactt cccgggggctc	240
aggctcaggg acttctgggt gtagtggttg tgcagggcct cgtgcatcac gctgcagctg	300
aacacgttgc cctgctgcca ccggctcttg tccacggtca gcttgctata caggaagaat	360
gagccgtcgc tgtccagcac aggggggggtg gtcttgtagt tgttctcggg ctggccgttg	420
ctctcccatt ccacggcgat ctgctgggg tagaagccct tgaccaggca ggtcaggagc	480
acctggttct tggatcatctc ttcccggctg gggggcagtg tgtagacctg aggctcgcgg	540
ggctggccct tggccttgct gatggttttc tcgatggggg caggcagggc cttggttgac	600
accttgcaact tgtactcttt gccgttcagc cagtcctggt gcagcacggt cagcacggac	660
accacccggt aggtgctggt gtactgttcc tctctgggct tggctcttggc gttgtgcaact	720
tccacgccgt ccacgtacca attgaacttc acttcagggc cctcgtggga cagctccacc	780
accacgcagg tcaacttcggg ggtccggctg atcatcaggg tgccttggg ctttgggggg	840
aacaggaaca cgctggggcc tcccagcagt tcaggggcag ggcagggggg acacgtgtgg	900
gtcttgctgc agctcttagg ttccaccgc ttgtccacct tgggtgttgct gggcttgtgg	960
ttcacgttgc agatgtaggt ctgggtgccc aggctgctgc tgggcacggt gaccacgctg	1020
ctcaggctat acaggccgct gctctgcagc acggctggaa aggtgtgcac gccgctggtc	1080

agggcgccag agttccagga cacggtcacg ggctcgggga agtagtcctt gaccaggcag	1140
cccagggcgg ctgttccgcc agaggtgctc ttgctgctag gggccagagg gaacacgctt	1200
ggtccttgg tgctggcgct cgagacggtc accatggttc cctgtcccca gatatcgaag	1260
gttcctccct ccaggatatc catggcgctg ctctccagcc gatccttggc gcagtagtac	1320
agggcggtat cctcggcccg caggctgttc atctgcaggt acaggctgtt ctttccgtta	1380
tcccggctga tgggtgaaccg tccctgcacg ctatcggcgt atcccatgaa gtttcccttc	1440
cagttgattc cggccaccca ctccagtccc tttcctgggg cctgccgcac ccagtgcag	1500
gtgtactcat cgaagctgaa tccgctggcg gcgcagctca gccgcaggct ccgtcctggc	1560
tgcaccagtc ctctccgggt ctccaccagc tgcacctctg aattcgtcac cagggccagg	1620
ctcagggcga tcagcagcag cagctgcagc cgcatggctg cggcgcgatc tgacggttca	1680
ctaaacgagc tctgcttata taggcctccc accgtacacg ccacctcgac atacctagtt	1740
attaatagta atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgtta	1800
cataacttac ggtaaattggc ccgcctggct gaccgcccga cgacccccgc ccattgacgt	1860
caataatgac gtatgttccc atagtaacgc caatagggac tttccattga cgtcaatggg	1920
tggagtattht acggtaaaact gcccaattgg cagtacatca agtgtatcat atgccaagta	1980
cgccccctat tgacgtcaat gacggtaaat ggcccgctg gcattatgcc cagtacatga	2040
ccttatggga ctttctact tggcagtaca tctacgtatt agtcatcgct attaccatgg	2100
tgatgcgggt ttggcagtag atcaatgggc gtggatagcg gtttgactca cggggatttc	2160
caagtctcca cccattgac gtcaatggga gtttggtttg gcacaaaaat caacgggact	2220
ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat gggcggtagg cgtgtacgg	2280
gggaggtcta tataagcaga gctggtttag tgaaccgtca gatccgctgg gcactttgca	2340
ctggaactta caacaccgga gcaaggacgc gactctgccg cccaccatg cgcagtcagc	2400
tgctgctgct gatcgccctg agcctggccc tggtagacaa cagcgatatc gtcagaccc	2460
agagcccaga tagcctggcc gtgagcctgg gagagcgggc caccatcaac tgcaagagca	2520
gccagagcgt gctgtacagc agcaacaaca agaactacct ggccctggtag cagcagaagc	2580
caggacagcc accaaagctg ctgatctact gggccagcac ccgggagagc ggagtgccag	2640
atcggttcag cggaagcgga agcggaaaccg atttcacctt gaccatcagc agcctgcagg	2700

ccgaggatgt ggccgtgtac tactgccagc agtactacag caccctactg accttcggac 2760

agggaaccaa ggtggagatc aag cgt acg gtg gcc gcc cca agc gtg ttc atc 2813
Arg Thr Val Ala Ala Pro Ser Val Phe Ile
1 5 10

ttc cca cca agc gat gag cag ctg aag agc gga acc gcc agc gtg gtg 2861
Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val
15 20 25

tgc ctg ctg aac aac ttc tac cca cgg gag gcc aag gtg cag tgg aag 2909
Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys
30 35 40

gtg gat aac gcc ctg cag agc gga aac agc cag gag agc gtg acc gag 2957
Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu
45 50 55

cag gat agc aag gat agc acc tac agc ctg agc agc acc ctg acc ctg 3005
Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu
60 65 70

agc aag gcc gat tac gag aag cac aag gtg tac gcc tgc gag gtg acc 3053
Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr
75 80 85 90

cac cag gga ctg agc agc cca gtg acc aag agc ttc aac cgc gga gag 3101
His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu
95 100 105

tgc cggaagcggc gggccccagt gaagcagacc ctgaacttcg atctgctgaa 3154
Cys

gctggccgga gatgtggaga gcaaccacagg accaatgtac agaatgcagc tgctgagctg 3214

catcgccctg agcctggccc tggtagaccaa cagc cag gtg caa cta gtg gag agc 3269
Gln Val Gln Leu Val Glu Ser
110

gga gga gga gtg gtg cag cca gga cgg agc ctg cgg ctg agc tgc gcc 3317
Gly Gly Gly Val Val Gln Pro Gly Arg Ser Leu Arg Leu Ser Cys Ala
115 120 125 130

gcc agc gga ttc acc ttc agc acc tac gcc atg cac tgg gtg cgg cag 3365
Ala Ser Gly Phe Thr Phe Ser Thr Tyr Ala Met His Trp Val Arg Gln
135 140 145

gcc cca gga aag gga ctg gag tgg gtg gcc gtg atc agc tac gat gcc 3413
Ala Pro Gly Lys Gly Leu Glu Trp Val Ala Val Ile Ser Tyr Asp Ala
150 155 160

aac tac aag tac tac gcc gat agc gtg aag gga cgg ttc acc atc agc 3461
Asn Tyr Lys Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser

165	170	175	
cgg gat aac agc aag aac acc ctg tac ctg cag atg aac agc ctg cgg Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg 180 185 190			3509
gcc gag gat acc gcc gtg tac tac tgc gcc aag gat agc cag ctg cgg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Lys Asp Ser Gln Leu Arg 195 200 205 210			3557
agc ctg ctg tac ttc gag tgg ctg agc cag gga tac ttc gat tac tgg Ser Leu Leu Tyr Phe Glu Trp Leu Ser Gln Gly Tyr Phe Asp Tyr Trp 215 220 225			3605
gga cag gga acc ctg gtg acc gtg agc agc gct agc acc aag gga cca Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro 230 235 240			3653
agc gtg ttc cca ctg gcc cca agc agc aag agc acc agc gga gga acc Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr 245 250 255			3701
gcc gcc ctg gga tgc ctg gtg aag gat tac ttc cca gag cca gtg acc Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr 260 265 270			3749
gtg agc tgg aac agc gga gcc ctg acc agc gga gtg cac acc ttc cca Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro 275 280 285 290			3797
gcc gtg ctg cag agc agc gga ctg tat agc ctg agc agc gtg gtg acc Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr 295 300 305			3845
gtg cca agc agc agc ctg gga acc cag acc tac atc tgc aac gtg aac Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn 310 315 320			3893
cac aag cca agc aac acc aag gtg gat aag aag gtg gag cca aag agc His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser 325 330 335			3941
tgc gat aag acc cac acg tgc cct cct tgt cca gcc ccc gaa ctg ctg Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu 340 345 350			3989
ggc ggg cct agc gtg ttc ctg ttt ccc cct aag cct aaa gat aca ctg Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu 355 360 365 370			4037
atg att agt aga acc cca gag gtc aca tgc gtg gtc gtg gac gtg tcc Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser 375 380 385			4085

cac gaa gag cct gac gtg aag ttc aac tgg tac gtg gat ggc gtg gag His Glu Glu Pro Asp Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu 390 395 400	4133
gtg cac aat gct aag act aaa cca cgc gaa gag cag tat aat agt aca Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr 405 410 415	4181
tac cga gtc gtg tca gtc ctg aca gtg ctg cac cag gat tgg ctg aac Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn 420 425 430	4229
ggc aag gag tat aag tgc aag gtg tct aac aag gcc ctg ccc gcc cct Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro 435 440 445 450	4277
atc gag aaa aca att agc aag gcc aaa ggg cag cca cgg gaa ccc cag Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln 455 460 465	4325
gtc tac act ctg cca ccc tca aga gat gaa ctg act aag aac cag gtc Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val 470 475 480	4373
agc ctg acc tgt ctg gtg aaa ggc ttc tac ccc agc gac atc gcc gtg Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val 485 490 495	4421
gag tgg gaa agt aac ggc cag cct gag aat aac tac aag act acc cct Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro 500 505 510	4469
cca gtg ctg gat agc gac ggg tcc ttc ttc ctg tat agc aag ctg aca Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr 515 520 525 530	4517
gtg gac aaa tcc cgc tgg cag cag gga aac gtc ttt tcc tgt tct gtg Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val 535 540 545	4565
atg cat gag gcc ctg cac aat cat tac acc cag aag agt ctg tca ctg Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu 550 555 560	4613
agc ccc ggc aaa tgataaaagg aaccgcgcgt atgacggcaa taaaaagaca Ser Pro Gly Lys 565	4665
gaataaaacc cacgggtggtt gggtcgtttg ttcataaacc cgggatcgat aaggatcttc	4725
ctagagcatg gctacgtaga taagtagcat ggcggggttaa tcattaacta caaggaaccc	4785
ctagtgatgg agttggccac tccctctctg cgcgctcgct cgctcactga ggccggggcga	4845

ccaaaggtcg	cccgaacgcc	gggctttgcc	cgggcggcct	cagtgaacga	gcgagcgcgc	4905
agccttaatt	aacctaatc	actggccgct	gttttacaac	gtcgtgactg	ggaaaaccct	4965
ggcgttaccc	aacttaatcg	ccttgacgca	catccccctt	tcgccagctg	gcgtaatagc	5025
gaagaggccc	gcaccgatcg	cccttcccaa	cagttgcgca	gcctgaatgg	cgaatgggac	5085
gcgccctgta	gcggcgcat	aagcgcggcg	ggtgtggtgg	ttacgcgcag	cgtgaccgct	5145
acacttgcca	gcgccttagc	gcccgctcct	ttcgctttct	tcccttcctt	tctcgccacg	5205
ttcgccggct	ttccccgtca	agctctaaat	cgggggctcc	ctttaggggt	cggatttagt	5265
gctttacggc	acctcgaccc	caaaaaactt	gattaggggt	atggttcacg	tagtgggcca	5325
tcgccctgat	agacggtttt	tcgccctttg	acgttgaggt	ccacgttctt	taatagtgga	5385
ctcttgttcc	aaactggaac	aacactcaac	cctatctcgg	tctattcttt	tgatttataa	5445
gggattttgc	cgatttcggc	ctattggtta	aaaaatgagc	tgatttaaca	aaaatttaac	5505
gcgaatttta	acaaaatatt	aacgcttaca	atttaggtgg	cacttttcgg	ggaaatgtgc	5565
gcggaacccc	tatttgttta	tttttctaaa	tacattcaaa	tatgtatccg	ctcatgagac	5625
aataaccctg	ataaatgctt	caataatatt	gaaaaaggaa	gagtatgagt	attcaacatt	5685
tccgtgtcgc	ccttattccc	ttttttgcgg	cattttgcct	tcctgttttt	gctcaccacg	5745
aaacgctggt	gaaagtaaaa	gatgctgaag	atcagttggg	tgacagagtg	ggttacatcg	5805
aactggatct	caacagcggg	aagatccttg	agagttttcg	ccccgaagaa	cgttttccaa	5865
tgatgagcac	ttttaaagtt	ctgctatgtg	gcgcggtatt	atcccgat	gacgcggggc	5925
aagagcaact	cggtcgcgcg	atacactatt	ctcagaatga	cttggttgag	tactcaccag	5985
tcacagaaaa	gcattcttacg	gatggcatga	cagtaagaga	attatgcagt	gctgccataa	6045
ccatgagtga	taacactgcg	gccaaacttac	ttctgacaac	gatcggagga	cgaaggagc	6105
taaccgcttt	tttgacacaac	atgggggatc	atgtaactcg	ccttgatcgt	tggaaccgg	6165
agctgaatga	agccatacca	aacgacgagc	gtgacaccac	gatgcctgta	gcaatggcaa	6225
caacgttgcg	caaaactatta	actggcgaac	tacttactct	agcttcccgg	caacaattaa	6285
tagactggat	ggaggcggat	aaagttgcag	gaccacttct	gcgctcggcc	cttcgggctg	6345
gctggtttat	tgctgataaa	tctggagccg	gtgagcgtgg	gtctcgcggg	atcattgcag	6405
cactggggcc	agatggtaag	ccctcccgta	tcgtagtatt	ctacacgacg	gggagtcagg	6465

caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt 6525
 ggtaactgtc agaccaagtt tactcatata tactttagat tgatttaaaa cttcattttt 6585
 aatttaaaag gatctaggtg aagatccttt ttgataatct catgaccaa atcccttaac 6645
 gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag 6705
 atcctttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg 6765
 tggtttggtt gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca 6825
 gagcgcagat accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga 6885
 actctgtagc accgcctaca tacctcgctc tgctaatacct gttaccagtg gctgctgcca 6945
 gtggcgataa gtcgtgtctt accgggttg actcaagacg atagttaccg gataaggcgc 7005
 agcggtcggg ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca 7065
 ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc gaaggagaaa 7125
 aggcggacag gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc 7185
 cagggggaaa cgcttggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc 7245
 gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg 7305
 cttttttacg gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat 7365
 cccctgattc tgtggataac cgtattaccg cttttgagtg agctgatacc gctcgccgca 7425
 gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca 7485
 aaccgcctct ccccgcgctg tggccgattc attaatgcag ctggcacgac aggtttcccg 7545
 actggaaagc gggcagtgag cgcaacgcaa ttaatgtgag ttagctcact cattaggcac 7605
 cccaggcttt acactttatg cttccggctc gtatgttggtg tggaattgtg agcggataac 7665
 aatttcacac aggaaacagc tatgaccatg attacgccag atttaattaa ggccttaatt 7725
 aggctgcgcg ctgcctcgct c 7746

<210> 20
 <211> 107
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 20

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
1 5 10 15

Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe
20 25 30

Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln
35 40 45

Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
50 55 60

Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
65 70 75 80

Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
85 90 95

Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
100 105

<210> 21
<211> 129
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc tổng hợp

<400> 21

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr
20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ala Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu Ser
100 105 110

Gln Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
115 120 125

Ser

<210> 22
<211> 107
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc tổng hợp

<400> 22

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
65 70 75 80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
85 90 95

Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His
 100 105

<210> 23
 <211> 223
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 23

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
 1 5 10 15

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
 20 25 30

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Glu Pro Asp
 35 40 45

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
 50 55 60

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

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
 85 90 95

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
 100 105 110

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
 115 120 125

Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
 130 135 140

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
 145 150 155 160

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
 165 170 175

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
 180 185 190

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
 195 200 205

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 210 215 220

<210> 24
 <211> 7740
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Plasmit mang kháng thể đơn dòng FI6 và CR8033

<220>
 <221> tín hiệu_polyA
 <222> (191)..(239)
 <223> tổng hợp polyA

<220>
 <221> đặc điểm_đa dạng
 <222> (246)..(914)
 <223> bổ sung- CH'2-3

<220>
 <221> đặc điểm_đa dạng
 <222> (915)..(1235)
 <223> bổ sung- CH'1

<220>
 <221> đặc điểm_đa dạng
 <222> (1236)..(1622)
 <223> bổ sung- CR8033\ VH

<220>
 <221> đặc điểm_đa dạng
 <222> (1623)..(1673)
 <223> bổ sung- đoạn dẫn

<220>
 <221> đặc điểm_đa dạng
 <222> (1683)..(1751)

```

<223> CMV\mp2

<220>
<221> đặc điểm_đa dạng
<222> (1752)..(2220)
<223> Đoạn tăng cường

<220>
<221> đặc điểm_đa dạng
<222> (2221)..(2346)
<223> CMV\mp1

<220>
<221> đặc điểm_đa dạng
<222> (2406)..(2462)
<223> đoạn dẫn

<220>
<221> CDS
<222> (2463)..(2795)
<223> FI6\VL

<220>
<221> CDS
<222> (2796)..(3116)
<223> CL

<220>
<221> đặc điểm_đa dạng
<222> (3201)..(3260)
<223> đoạn dẫn

<220>
<221> CDS
<222> (3261)..(3647)
<223> FI6\ VH

<220>
<221> CDS
<222> (3648)..(3968)
<223> CH1

<220>
<221> CDS
<222> (3969)..(4637)
<223> CH2-3

<220>
<221> đặc điểm_đa dạng
<222> (3980)..(3980)
<223> A -> T

<220>
<221> tín hiệu_polyA

```

<222> (4644)..(4721)

<223> TKpA ngắn

<400> 24

ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgctcg ggcgaccttt	60
ggtcgccccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact	120
aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct	180
aggaagatct cacacaaaaa accaacacac agatgtaatg aaaataaaga tattttatatt	240
tatcacttcc cggggctcag gctcagggac ttctgggtgt agtggttgtg cagggcctcg	300
tgcatacgc tgcagctgaa cacgttgccc tgctgccacc ggctcttgtc cacggtcagc	360
ttgctataca ggaagaatga gccgtcgctg tccagcacag ggggggtggt cttgtagttg	420
ttctcgggct ggccgttgct ctcccattcc acggcgatct cgctggggta gaagcccttg	480
accaggcagg tcagggacac ctggttcttg gtcattctct cccggctggg gggcagtgtg	540
tagacctgag gctcgcgggg ctggcccttg gccttgctga tggttttctc gatgggggca	600
ggcagggcct tgttggaacac cttgcacttg tactctttgc cgttcagcca gtccctggtgc	660
agcacggtca gcacggacac cacccgtag gtgctgttgt actgttcctc tctgggcttg	720
gtcttggcgt tgtgcacttc cacgccgtcc acgtaccaat tgaacttcac ttcagggtcc	780
tcgtgggaca cgtccaccac cacgcaggtc acttcggggg tccggctgat catcagggtg	840
tccttgggct ttggggggaa caggaacacg ctggggcctc ccagcagttc aggggcaggg	900
caggggggac acgtgtgggt cttgtcgag ctcttaggtt ccaccgctt gtccaccttg	960
gtgttgctgg gcttggtggt cacgttgag atgtaggtct gggtgcccag gctgctgctg	1020
ggcacggtga ccacgtgct caggctatac aggccgctgc tctgcagcac ggctggaaaag	1080
gtgtgcacgc cgctggtcag ggcgccagag ttccaggaca cggtcacggg ctcggggaag	1140
tagtccttga ccaggcagcc cagggcggct gttccgccag aggtgctctt gctgctaggg	1200
gccagagggg acacgcttgg tcccttggtg ctggcgctcg agacggtcac catggttccc	1260
tgtccccaga tatcgaaggt tcctccctcc aggatatcca tggcgctgct ctccagccga	1320
tccttggcgc agtagtacag ggcggtatcc tcggcccga ggctgttcat ctgcaggtag	1380
aggctgttct ttccgttatc ccggctgatg gtgaaccgtc cctgcacgct atcggcgtat	1440
cccatgaagt ttcccttcca gttgattccg gccaccact ccagtccctt tcctggggcc	1500

tgccgcaccc agtgcattggt gtactcatcg aagctgaatc cgctggcggc gcagctcagc	1560
cgcaggctcc gtccctggctg caccagtcct cctccggtct ccaccagctg cacctctgaa	1620
ttcgtcacca gggccaggct cagggcgatc agcagcagca gctgcatgcg catggtggcg	1680
gcgcgatctg acggttcact aaacgagctc tgcttatata ggccctccac cgtacacgcc	1740
acctcgacat acctagttat taatagtaat caattacggg gtcattagtt catagcccat	1800
atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga ccgcccacg	1860
acccccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt	1920
tccattgacg tcaatgggtg gagtatctac ggtaaactgc ccacttgga gtacatcaag	1980
tgtatcatat gccaagtacg cccctattg acgtcaatga cggtaaattg cccgcctggc	2040
attatgccca gtacatgacc ttatgggact ttcctacttg gcagtacatc tacgtattag	2100
tcacgcctat taccatgggtg atgcggtttt ggagctacat caatgggcgt ggatagcggt	2160
ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt ttgttttggc	2220
acaaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg acgcaaatgg	2280
gcggtagggc tgtacgggtg gaggtctata taagcagagc tggtttagtg aaccgtcaga	2340
tccgctgggc actttgcact ggaacttaca acacccgagc aaggacgca ctctgccgcc	2400
ccaccatgcg catgcagctg ctgctgctga tcgccctgag cctggccctg gtgaccaaca	2460
gc gat atc gtc atg acc cag agc cca gat agc ctg gcc gtg agc ctg	2507
Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu	
1 5 10 15	
gga gag cgg gcc acc atc aac tgc aag agc agc cag agc gtg acc ttc	2555
Gly Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Thr Phe	
20 25 30	
aac tac aag aac tac ctg gcc tgg tac cag cag aag cca gga cag cca	2603
Asn Tyr Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Pro	
35 40 45	
cca aag ctg ctg atc tac tgg gcc agc acc cgg gag agc gga gtg cca	2651
Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro	
50 55 60	
gat cgg ttc agc gga agc gga agc gga acc gat ttc acc ctg acc atc	2699
Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile	
65 70 75	
agc agc ctg cag gcc gag gat gtg gcc gtg tac tac tgc cag cag cac	2747
Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln His	

80	85	90	95	
tac cgg acc cca cca acc ttc gga cag gga acc aag gtg gag atc aag				2795
Tyr Arg Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys	100	105	110	
cgt acg gtg gcc gcc cca agc gtg ttc atc ttc cca cca agc gat gag				2843
Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu	115	120	125	
cag ctg aag agc gga acc gcc agc gtg gtg tgc ctg ctg aac aac ttc				2891
Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe	130	135	140	
tac cca cgg gag gcc aag gtg cag tgg aag gtg gat aac gcc ctg cag				2939
Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln	145	150	155	
agc gga aac agc cag gag agc gtg acc gag cag gat agc aag gat agc				2987
Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser	160	165	170	175
acc tac agc ctg agc agc acc ctg acc ctg agc aag gcc gat tac gag				3035
Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu	180	185	190	
aag cac aag gtg tac gcc tgc gag gtg acc cac cag gga ctg agc agc				3083
Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser	195	200	205	
cca gtg acc aag agc ttc aac cgc gga gag tgc cggaagcggc gggccccagt				3136
Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys	210	215		
gaagcagacc ctgaacttcg atctgctgaa gctggccgga gatgtggaga gcaacccagg				3196
accaatgtac agaatgcagc tgctgagctg catcgccctg agcctggccc tggtgaccaa				3256
cagc cag gtg caa cta gtg gag agc gga gga gga gtg gtg cag cca gga				3305
Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly	220	225	230	
cgg agc ctg cgg ctg agc tgc gcc gcc agc gga ttc acc ttc agc acc				3353
Arg Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr	235	240	245	
tac gcc atg cac tgg gtg cgg cag gcc cca gga aag gga ctg gag tgg				3401
Tyr Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp	250	255	260	265
gtg gcc gtg atc agc tac gat gcc aac tac aag tac tac gcc gat agc				3449
Val Ala Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser	270	275	280	

gtg aag gga cgg ttc acc atc agc cgg gat aac agc aag aac acc ctg	3497
Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu	
285 290 295	
tac ctg cag atg aac agc ctg cgg gcc gag gat acc gcc gtg tac tac	3545
Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr	
300 305 310	
tgc gcc aag gat agc cag ctg cgg agc ctg ctg tac ttc gag tgg ctg	3593
Cys Ala Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu	
315 320 325	
agc cag gga tac ttc gat tac tgg gga cag gga acc ctg gtg acc gtg	3641
Ser Gln Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val	
330 335 340 345	
agc agc gct agc acc aag gga cca agc gtg ttc cca ctg gcc cca agc	3689
Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser	
350 355 360	
agc aag agc acc agc gga gga acc gcc gcc ctg gga tgc ctg gtg aag	3737
Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys	
365 370 375	
gat tac ttc cca gag cca gtg acc gtg agc tgg aac agc gga gcc ctg	3785
Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu	
380 385 390	
acc agc gga gtg cac acc ttc cca gcc gtg ctg cag agc agc gga ctg	3833
Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu	
395 400 405	
tat agc ctg agc agc gtg gtg acc gtg cca agc agc agc ctg gga acc	3881
Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr	
410 415 420 425	
cag acc tac atc tgc aac gtg aac cac aag cca agc aac acc aag gtg	3929
Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val	
430 435 440	
gat aag aag gtg gag cca aag agc tgc gat aag acc cac acg tgc cct	3977
Asp Lys Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro	
445 450 455	
cct tgt cca gcc ccc gaa ctg ctg ggc ggg cct agc gtg ttc ctg ttt	4025
Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe	
460 465 470	
ccc cct aag cct aaa gat aca ctg atg att agt aga acc cca gag gtc	4073
Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val	
475 480 485	
aca tgc gtg gtc gtg gac gtg tcc cac gaa gag cct gac gtg aag ttc	4121
Thr Cys Val Val Val Asp Val Ser His Glu Glu Pro Asp Val Lys Phe	

490	495	500	505	
aac tgg tac gtg gat ggc gtg gag gtg cac aat gct aag act aaa cca				4169
Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro	510	515	520	
cgc gaa gag cag tat aat agt aca tac cga gtc gtg tca gtc ctg aca				4217
Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr	525	530	535	
gtg ctg cac cag gat tgg ctg aac ggc aag gag tat aag tgc aag gtg				4265
Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val	540	545	550	
tct aac aag gcc ctg ccc gcc cct atc gag aaa aca att agc aag gcc				4313
Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala	555	560	565	
aaa ggg cag cca cgg gaa ccc cag gtc tac act ctg cca ccc tca aga				4361
Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg	570	575	580	585
gat gaa ctg act aag aac cag gtc agc ctg acc tgt ctg gtg aaa ggc				4409
Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly	590	595	600	
ttc tac ccc agc gac atc gcc gtg gag tgg gaa agt aac ggc cag cct				4457
Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro	605	610	615	
gag aat aac tac aag act acc cct cca gtg ctg gat agc gac ggg tcc				4505
Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser	620	625	630	
ttc ttc ctg tat agc aag ctg aca gtg gac aaa tcc cgc tgg cag cag				4553
Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln	635	640	645	
gga aac gtc ttt tcc tgt tct gtg atg cat gag gcc ctg cac aat cat				4601
Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His	650	655	660	665
tac acc cag aag agt ctg tca ctg agc ccc ggc aaa tgataaaagg				4647
Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys	670	675		
aacccgcgct atgacggcaa taaaaagaca gaataaaacc cacgggtggtt gggtcgtttg				4707
ttcataaacc cgggatcgat aaggatcttc ctagagcatg gctacgtaga taagtagcat				4767
ggcggggttaa tcattaacta caaggaaccc ctagtgatgg agttggccac tccctctctg				4827
cgcgctcgct cgctcactga ggccgggcca ccaaaggctcg cccgacgccc gggctttgcc				4887

cgggcggcct cagtgagcga gcgagcgcgc agccttaatt aacctaattc actggcgcgc	4947
gttttacaac gtcgtgactg ggaaaaccct ggcgttaccc aacttaatcg ccttgcagca	5007
catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa	5067
cagttgcgca gcctgaatgg cgaatgggac gcgccctgta gcggcgcatt aagcgcggcg	5127
ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct	5187
ttcgctttct tcccttcctt tctcgccacg ttcgcccggct ttccccgtca agctctaaat	5247
cgggggctcc ctttaggggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt	5307
gattaggggtg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg	5367
acgttgagtg ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac	5427
cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta	5487
aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca	5547
atthaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa	5607
tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt	5667
gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg	5727
cattttgcct tcctgttttt gctcaccag aaacgctggg gaaagtaaaa gatgctgaag	5787
atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggg aagatccttg	5847
agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg	5907
gcgcggtatt atcccgatt gacgccgggc aagagcaact cggtcgccgc atacactatt	5967
ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga	6027
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaaacttac	6087
ttctgacaac gatcggagga ccgaaggagc taaccgcttt ttgacacaac atgggggatc	6147
atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc	6207
gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac	6267
tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag	6327
gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg	6387
gtgagcgtgg gtctcgcggg atcattgcag cactggggcc agatggtaag ccctcccgta	6447
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg	6507

ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata 6567
 tacttttagat tgatttataaa cttcattttt aatttataaag gatctaggtg aagatccttt 6627
 ttgataatct catgacaaaa atcccttaac gtgagttttc gttccactga gcgtcagacc 6687
 ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct 6747
 tgcaaacaaa aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa gagctaccaa 6807
 ctctttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gttcttctag 6867
 tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc 6927
 tgctaatacct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accggggttg 6987
 actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca 7047
 cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat 7107
 gagaaagcgc cacgcttccc gaaggagaaa aggcggacag gtatccggta agcggcaggg 7167
 tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc 7227
 ctgtcggggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc 7287
 ggagcctatg gaaaaacgcc agcaacgcgg cttttttacg gttcctggcc ttttgctggc 7347
 cttttgctca catgttcttt cctgcgttat ccctgattc tgtggataac cgtattaccg 7407
 cttttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga 7467
 gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgcggt tggccgattc 7527
 attaattgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa 7587
 ttaatgtgag ttagctcact cattaggcac ccaggcttt acactttatg cttccggctc 7647
 gtatgttggtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg 7707
 attacgccag atttaattaa ggccttaatt agg 7740

<210> 25
 <211> 111
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 25

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly

1				5					10					15			
Glu	Arg	Ala	Thr	Ile	Asn	Cys	Lys	Ser	Ser	Gln	Ser	Val	Thr	Phe	Asn		
			20					25					30				
Tyr	Lys	Asn	Tyr	Leu	Ala	Trp	Tyr	Gln	Gln	Lys	Pro	Gly	Gln	Pro	Pro		
		35					40					45					
Lys	Leu	Leu	Ile	Tyr	Trp	Ala	Ser	Thr	Arg	Glu	Ser	Gly	Val	Pro	Asp		
	50					55					60						
Arg	Phe	Ser	Gly	Ser	Gly	Ser	Gly	Thr	Asp	Phe	Thr	Leu	Thr	Ile	Ser		
65					70					75					80		
Ser	Leu	Gln	Ala	Glu	Asp	Val	Ala	Val	Tyr	Tyr	Cys	Gln	Gln	His	Tyr		
				85					90					95			
Arg	Thr	Pro	Pro	Thr	Phe	Gly	Gln	Gly	Thr	Lys	Val	Glu	Ile	Lys			
			100					105					110				
<210> 26																	
<211> 107																	
<212> PRT																	
<213> Trình tự nhân tạo																	
<220>																	
<223> Cấu trúc tổng hợp																	
<400> 26																	
Arg	Thr	Val	Ala	Ala	Pro	Ser	Val	Phe	Ile	Phe	Pro	Pro	Ser	Asp	Glu		
1				5					10					15			
Gln	Leu	Lys	Ser	Gly	Thr	Ala	Ser	Val	Val	Cys	Leu	Leu	Asn	Asn	Phe		
			20					25					30				
Tyr	Pro	Arg	Glu	Ala	Lys	Val	Gln	Trp	Lys	Val	Asp	Asn	Ala	Leu	Gln		
		35					40					45					
Ser	Gly	Asn	Ser	Gln	Glu	Ser	Val	Thr	Glu	Gln	Asp	Ser	Lys	Asp	Ser		
	50					55					60						
Thr	Tyr	Ser	Leu	Ser	Ser	Thr	Leu	Thr	Leu	Ser	Lys	Ala	Asp	Tyr	Glu		

[illegible]

<210> 28
 <211> 107
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 28

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
 1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
 20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
 35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
 50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
 65 70 75 80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
 85 90 95

Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His
 100 105

<210> 29
 <211> 223
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 29

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
 1 5 10 15

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
20 25 30

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Glu Pro Asp
35 40 45

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
50 55 60

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

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
85 90 95

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
100 105 110

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
115 120 125

Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
130 135 140

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
145 150 155 160

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
165 170 175

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
180 185 190

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
195 200 205

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
210 215 220

<210> 30

<211> 7782
 <212> ADN
 <213> Nhân tạo

<220>
 <223> EcoRV

<220>
 <221> tín hiệu_polyA
 <222> (201)..(252)
 <223> bổ sung- tổng hợp\polyA

<220>
 <221> đặc điểm đa dạng
 <222> (268)..(588)
 <223> bổ sung- CL

<220>
 <221> đặc điểm đa dạng
 <222> (589)..(909)
 <223> bổ sung- TCN032\VL

<220>
 <221> tín hiệu_polyA
 <222> (910)..(966)
 <223> bổ sung- đoạn dẫn

<220>
 <221> đặc điểm đa dạng
 <222> (1026)..(1094)
 <223> bổ sung- CMV\mp2

<220>
 <221> đặc điểm đa dạng
 <222> (1095)..(1563)
 <223> Đoạn tăng cường

<220>
 <221> đặc điểm đa dạng
 <222> (1564)..(1689)
 <223> CMV\mp1

<220>
 <221> đặc điểm đa dạng
 <222> (1749)..(1805)
 <223> đoạn dẫn

<220>
 <221> CDS
 <222> (1806)..(2165)
 <223> TCN032\XH

<220>

<221> CDS
 <222> (2166)..(2459)
 <223> CH1

<220>
 <221> đặc điểm đa dạng
 <222> (2460)..(3152)
 <223> khớp nối-CH2'-CH3'

<220>
 <221> đặc điểm đa dạng
 <222> (3239)..(3296)
 <223> đoạn dẫn

<220>
 <221> CDS
 <222> (3297)..(3683)
 <223> FI6\XH

<220>
 <221> CDS
 <222> (3684)..(4004)
 <223> CH1

<220>
 <221> CDS
 <222> (4005)..(4673)
 <223> CH2-3

<220>
 <221> tín hiệu polyA
 <222> (4693)..(4770)
 <223> TKpA ngắn

<400> 30
 ggccttaatt aggctgcgcg ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg 60
 tcgggcgacc tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc 120
 caactccatc actagggggtt ccttgtagtt aatgattaac ccgcatgct acttatctac 180
 gtagccatgc tctaggaaga tctcacacaa aaaaccaaca cacagatgta atgaaaataa 240
 agatatTTTA ttgcggccgc tttatcagca ctctccgcgg ttgaagctct tggtcactgg 300
 gctgctcagt ccctgggtggg tcacctcgca ggcgtacacc ttgtgcttct cgtaatcggc 360
 cttgctcagg gtcagggtgc tgctcaggct gtaggtgcta tccttgctat cctgctcggt 420
 cacgctctcc tggctgtttc cgctctgcag ggcgttatcc acctccact gcaccttggc 480
 ctcccgtggg tagaagttgt tcagcaggca caccacgctg gcggttccgc tcttcagctg 540
 ctcatcgctt ggtgggaaga tgaacacgct tggggcgggc accgtacgct tgatctccac 600

```

ccgggttcct cctccgaagg tcagtgggtg gctgtagctc tgctggcagt agtaggtggc 660
gaaatcctct ggctgcaggc tggatgatgg cagggtgaaa tcggttccgc ttccgcttcc 720
gctgaaccgg cttggcactc cgctctgcag tccgctggcg gcgctgatca gtccctttgg 780
ggcctttcct ggccgctgct ggtaccagtt caggacttg tagatgttct ggctggcccg 840
gcaggtgatg gtcacccgat ctcccacgct ggcgctcagg ctgcttgggc tctgggtcat 900
ctggatatcg ctggttggtca ccagggccag gctcagggcg atcagcagca gcagctgcat 960
tctcatgggtg gagagtcgcg tccttgctcg ggtgttgtaa gttccagtgc aaagtgcccc 1020
aattggcgat ctgacgggttc actaaacgag ctctgcttat ataggcctcc caccgtacac 1080
gccacctcga catacctagt tattaatagt aatcaattac ggggtcatta gttcatagcc 1140
catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgcccc 1200
acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg ccaatagggg 1260
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc 1320
aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcc 1380
ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat 1440
tagtcatcgc tattaccatg gtgatgcggt tttggcagta catcaatggg cgtggatagc 1500
ggtttgactc acggggattt ccaagtctcc accccattga cgtcaatggg agtttgtttt 1560
ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa 1620
tgggcggtag gcgtgtacgg tgggaggtct atataagcag agctggttta gtgaaccgtc 1680
agatccgctg ctagegggca ctttgactg gaacttaca caccgagca aggacgcgac 1740
tctccaccat gcgcatgcag ctgctgctgc tgatcgccct gagcctggcc ctggtgacca 1800
acagc cag gtg cag ctg cag gag agc gga cca gga ctg gtg aag cca agc 1850
      Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser
      1           5           10           15

gag acc ctg agc ctg acc tgc acc gtg agc gga agc agc atc agc aac 1898
Glu Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Ser Ser Ile Ser Asn
      20           25           30

tac tac tgg agc tgg atc cgg cag agc cca gga aag gga ctg gag tgg 1946
Tyr Tyr Trp Ser Trp Ile Arg Gln Ser Pro Gly Lys Gly Leu Glu Trp
      35           40           45

atc gga ttc atc tac tac gga gga aac acc aag tac aac cca agc ctg 1994

```

Ile	Gly	Phe	Ile	Tyr	Tyr	Gly	Gly	Asn	Thr	Lys	Tyr	Asn	Pro	Ser	Leu		
		50					55					60					
aag	agc	cgg	gtg	acc	atc	agc	cag	gat	acc	agc	aag	agc	cag	gtg	agc		2042
Lys	Ser	Arg	Val	Thr	Ile	Ser	Gln	Asp	Thr	Ser	Lys	Ser	Gln	Val	Ser		
	65					70					75						
ctg	acc	atg	agc	agc	gtg	acc	gcc	gcc	gag	agc	gcc	gtg	tac	ttc	tgc		2090
Leu	Thr	Met	Ser	Ser	Val	Thr	Ala	Ala	Glu	Ser	Ala	Val	Tyr	Phe	Cys		
80					85					90					95		
gcc	cgg	gcc	agc	tgc	agc	gga	gga	tac	tgc	atc	ctg	gat	tac	tgg	gga		2138
Ala	Arg	Ala	Ser	Cys	Ser	Gly	Gly	Tyr	Cys	Ile	Leu	Asp	Tyr	Trp	Gly		
				100					105					110			
cag	gga	acc	ctg	gtg	acc	gtg	agc	agc	gcg	tcg	acc	aag	gga	cct	tcg		2186
Gln	Gly	Thr	Leu	Val	Thr	Val	Ser	Ser	Ala	Ser	Thr	Lys	Gly	Pro	Ser		
			115					120					125				
gtc	ttc	ccc	ctg	gca	ccc	tcc	tcc	aag	agc	acc	tct	ggg	ggc	aca	gcg		2234
Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys	Ser	Thr	Ser	Gly	Gly	Thr	Ala		
		130					135					140					
gcc	ctg	ggc	tgc	ctg	gtc	aag	gac	tac	ttc	ccc	gaa	ccg	gtg	acg	gtg		2282
Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	Phe	Pro	Glu	Pro	Val	Thr	Val		
	145					150					155						
tcg	tgg	aac	tca	ggc	gcc	ctg	acc	agc	ggc	gtg	cac	acc	ttc	ccg	gct		2330
Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	Gly	Val	His	Thr	Phe	Pro	Ala		
160					165					170					175		
gtc	cta	cag	tcc	tca	gga	ctc	tac	tcc	ctc	agc	agc	gtg	gtg	acc	gtg		2378
Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val	Thr	Val		
				180					185					190			
ccc	tcc	agc	agc	ttg	ggc	acc	cag	acc	tac	atc	tgc	aac	gtg	aat	cac		2426
Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	Cys	Asn	Val	Asn	His		
			195					200					205				
aag	ccc	agc	aac	acc	aag	gtg	gac	aag	aaa	gtt	gaaccaaaga	gctgcgacaa					2479
Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val							
		210					215										
gacccacacg	tgtccccct	gccctgcccc	tgaactgctg	ggaggcccca	gcgtgttcct												2539
gttcccccca	aagcccaagg	acaccctgat	gatcagccgg	acccccgaag	tgacctgcgt												2599
ggtggtggac	gtgtcccacg	aggaccctga	agtgaagttt	aattggtacg	tggacggcgt												2659
ggaagtgcac	aacgccaaga	ccaagcccag	agaggaacag	tacaacagca	cctaccgggt												2719
ggtgtccgtg	ctgaccgtgc	tgcaccagga	ctggctgaac	ggcaaagagt	acaagtgcaa												2779
ggtgtccaac	aaggccctgc	ctgcccccat	cgagaaaacc	atcagcaagg	ccaagggcca												2839

```

gccccgcgag cctcaggtct acacactgcc ccccagccgg gaagagatga ccaagaacca      2899
ggtgtccctg acctgcctgg tcaagggctt ctaccccagc gacatcgccg tggaatggga      2959
gagcaacggc cagccccgaga acaactacaa gaccaccccc cctgtgctgg acagcgacgg      3019
ctcattcttc ctgtatagca agctgaccgt ggacaagagc cgggtggcagc agggcaacgt      3079
gttcagctgc agcgtgatgc acgaggccct gcacaaccac tacaccaga agtccctgag      3139
cctgagcccc ggcagaaagc ggagagcccc cgtgaagcag accctgaact tcgacctgct      3199
gaagctggcc ggcgacgtgg aaagcaaccc tggccctatg tacagaatgc agctgctgag      3259
ctgcatcgcc ctgagcctgg ccctggtgac caacagc cag gtg caa cta gtg gag      3314
                               Gln Val Gln Leu Val Glu
                               220

agc gga gga gga gtg gtg cag cca gga cgg agc ctg cgg ctg agc tgc      3362
Ser Gly Gly Gly Val Val Gln Pro Gly Arg Ser Leu Arg Leu Ser Cys
225                               230                               235                               240

gcc gcc agc gga ttc acc ttc agc acc tac gcc atg cac tgg gtg cgg      3410
Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr Ala Met His Trp Val Arg
245                               250                               255

cag gcc cca gga aag gga ctg gag tgg gtg gcc gtg atc agc tac gat      3458
Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ala Val Ile Ser Tyr Asp
260                               265                               270

gcc aac tac aag tac tac gcc gat agc gtg aag gga cgg ttc acc atc      3506
Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile
275                               280                               285

agc cgg gat aac agc aag aac acc ctg tac ctg cag atg aac agc ctg      3554
Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu
290                               295                               300

cgg gcc gag gat acc gcc gtg tac tac tgc gcc aag gat agc cag ctg      3602
Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Lys Asp Ser Gln Leu
305                               310                               315                               320

cgg agc ctg ctg tac ttc gag tgg ctg agc cag gga tac ttc gat tac      3650
Arg Ser Leu Leu Tyr Phe Glu Trp Leu Ser Gln Gly Tyr Phe Asp Tyr
325                               330                               335

tgg gga cag gga acc ctg gtg acc gtg agc agc gcc agc acc aag ggg      3698
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly
340                               345                               350

ccc agc gtg ttc cca ctg gcc cca agc agc aag agc acc agc gga gga      3746
Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly
355                               360                               365

```

acc gcc gcc ctg gga tgc ctg gtg aag gat tac ttc cca gag cca gtg	3794
Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val	
370 375 380	
acc gtg agc tgg aac agc gga gcc ctg acc agc gga gtg cac acc ttc	3842
Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe	
385 390 395 400	
cca gcc gtg ctg cag agc agc gga ctg tat agc ctg agc agc gtg gtg	3890
Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val	
405 410 415	
acc gtg cca agc agc agc ctg gga acc cag acc tac atc tgc aac gtg	3938
Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val	
420 425 430	
aac cac aag cca agc aac acc aag gtg gat aag aag gtg gag cca aag	3986
Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys	
435 440 445	
agc tgc gat aag acc cac acg tgc cct cca tgt cca gcc ccc gaa ctg	4034
Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu	
450 455 460	
ctg ggc ggg cct agc gtg ttc ctg ttt ccc cct aag cct aaa gat aca	4082
Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr	
465 470 475 480	
ctg atg att agt aga acc cca gag gtc aca tgc gtg gtc gtg gac gtg	4130
Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val	
485 490 495	
tcc cac gaa gag cct gac gtg aag ttc aac tgg tac gtg gat ggc gtg	4178
Ser His Glu Glu Pro Asp Val Lys Phe Asn Trp Tyr Val Asp Gly Val	
500 505 510	
gag gtg cac aat gct aag act aaa cca cgc gaa gag cag tat aat agt	4226
Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser	
515 520 525	
aca tac cga gtc gtg tca gtc ctg aca gtg ctg cac cag gat tgg ctg	4274
Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu	
530 535 540	
aac ggc aag gag tat aag tgc aag gtg tct aac aag gcc ctg ccc gcc	4322
Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala	
545 550 555 560	
cct atc gag aaa aca att agc aag gcc aaa ggg cag cca cgg gaa ccc	4370
Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro	
565 570 575	
cag gtc tac act ctg cca ccc tca aga gat gaa ctg act aag aac cag	4418

Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln	
580 585 590	
gtc agc ctg acc tgt ctg gtg aaa ggc ttc tac ccc agc gac atc gcc	4466
Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala	
595 600 605	
gtg gag tgg gaa agt aac ggc cag cct gag aat aac tac aag act acc	4514
Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr	
610 615 620	
cct cca gtg ctg gat agc gac ggg tcc ttc ttc ctg tat agc aag ctg	4562
Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu	
625 630 635 640	
aca gtg gac aaa tcc cgc tgg cag cag gga aac gtc ttt tcc tgt tct	4610
Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser	
645 650 655	
gtg atg cat gag gcc ctg cac aat cat tac acc cag aag agt ctg tca	4658
Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser	
660 665 670	
ctg agc ccc ggc aaa tgataaaaag cttctcgaga aggaacccgc gctatgacgg	4713
Leu Ser Pro Gly Lys	
675	
caataaaaag acagaataaa acccacgggt gttgggtcgt ttgttcataa acccggaag	4773
cttatcgata aggatcttcc tagagcatgg ctacgtagat aagtagcatg gcgggttaat	4833
cattaactac aaggaacccc tagtgatgga gttggccact ccctctctgc gcgctcgctc	4893
gctcactgag gccgggcgac caaaggctgc cgcacgccg ggctttgccc gggcggcctc	4953
agtgagcgag cgagcgcgca gccttaatta acctaattca ctggccgtcg ttttacaacg	5013
tcgtgactgg gaaaaccctg gcgttaccca acttaatcgc cttgcagcac atcccccttt	5073
cgccagctgg cgtaatagcg aagaggcccc caccgatcgc cttcccaac agttgcgcag	5133
cctgaatggc gaatgggacg cgccctgtag cggcgcatta agcgcggcgg gtgtgggtgt	5193
tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt	5253
cccttccttt ctgcacacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc	5313
tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg attaggggtga	5373
tggttcacgt agtgggcat cgccctgata gacggttttt cgcccttga cgttggagtc	5433
cacgttcttt aatagtggac tcttggtcca aactggaaca aactcaacc ctatctcggt	5493
ctattctttt gatttataag ggattttgcc gatttcggcc tattgggttaa aaaatgagct	5553

gatttaacaa aaattttaacg cgaatttttaa caaaatatta acgcttaciaa tttaggtggc	5613
acttttcggg gaaatgtgcg cggaacccct atttgtttat ttttctaaat acattcaa	5673
atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag	5733
agtatgagta ttcaacattt ccgtgtcgcc cttattccct tttttgcggc attttgcctt	5793
cctgtttttg ctcaccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt	5853
gcacgagtgg gttacatcga actggatctc aacagcggta agatccttga gagttttcgc	5913
cccgaagaac gttttccaat gatgagcact tttaaagtgc tgctatgtgg cgcggtatta	5973
tcccgtattg acgcgggca agagcaactc ggtcgccgca tacactattc tcagaatgac	6033
ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa	6093
ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg	6153
atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc	6213
cttgatcgtt gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg	6273
atgcctgtag caatggcaac aacgttgcg aaactattaa ctggcgaact acttactcta	6333
gcttcccggc aacaattaat agactggatg gaggcggata aagttgcagg accacttctg	6393
cgctcggccc ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg	6453
tctcgcggta tcattgcagc actggggcca gatggtaagc cctcccgtat cgtagttatc	6513
tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt	6573
gcctcactga ttaagcattg gtaactgtca gaccaagttt actcatatat acttttagatt	6633
gatttaaaac ttcattttta atttaaaagg atctaggtga agatcctttt tgataatctc	6693
atgacaaaaa tcccttaacg tgagttttcg ttccactgag cgtcagaccc cgtagaaaag	6753
atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa	6813
aaaccaccgc taccagcggg ggtttgtttg ccgatcaag agctaccaac tctttttccg	6873
aaggtaactg gcttcagcag agcgcagata ccaaatactg ttcttctagt gtagccgtag	6933
ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaactcctg	6993
ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga	7053
tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc	7113
ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc	7173

```

acgcttcccg aagggagaaa ggcggacagg tatccggtaa gcggcagggc cggaacagga 7233
gagcgcacga gggagcttcc agggggaaac gcctggtatc tttatagtcc tgcggggttt 7293
cgccacctct gacttgagcg tcgatttttg tgatgctcgt caggggggcg gagcctatgg 7353
aaaaacgcca gcaacgcggc ctttttacgg ttcctggcct tttgctggcc ttttgctcac 7413
atgttctttc ctgcgttatc ccctgattct gtggataacc gtattaccgc ctttgagtga 7473
gctgataccg ctgcgcgcag ccgaacgacc gagcgcagcg agtcagtgag cgaggaagcg 7533
gaagagcgcc caatacgcaa accgcctctc cccgcgcggt ggccgattca ttaatgcagc 7593
tggcacgaca ggtttcccgga ctggaaagcg ggcagtgagc gcaacgcaat taatgtgagt 7653
tagctcactc attaggcacc ccaggcttta cactttatgc ttccggctcg tatgttgtgt 7713
ggaattgtga gcggataaca atttcacaca ggaaacagct atgacatga ttacgccaga 7773
ttaattaa 7782

```

```

<210> 31
<211> 120
<212> PRT
<213> Nhân tạo

```

```

<220>
<223> Cấu trúc tổng hợp

```

```

<400> 31

```

```

Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu
1           5           10           15

```

```

Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Ser Ser Ile Ser Asn Tyr
20           25           30

```

```

Tyr Trp Ser Trp Ile Arg Gln Ser Pro Gly Lys Gly Leu Glu Trp Ile
35           40           45

```

```

Gly Phe Ile Tyr Tyr Gly Gly Asn Thr Lys Tyr Asn Pro Ser Leu Lys
50           55           60

```

```

Ser Arg Val Thr Ile Ser Gln Asp Thr Ser Lys Ser Gln Val Ser Leu
65           70           75           80

```

Thr Met Ser Ser Val Thr Ala Ala Glu Ser Ala Val Tyr Phe Cys Ala
85 90 95

Arg Ala Ser Cys Ser Gly Gly Tyr Cys Ile Leu Asp Tyr Trp Gly Gln
100 105 110

Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 32
<211> 98
<212> PRT
<213> Nhân tạo

<220>
<223> Cấu trúc tổng hợp

<400> 32

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
65 70 75 80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
85 90 95

Lys Val

<210> 33
<211> 129
<212> PRT
<213> Nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 33

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
 1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ala Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser Val
 50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
 65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95

Ala Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu Ser
 100 105 110

Gln Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 115 120 125

Ser

<210> 34

<211> 107

<212> PRT

<213> Nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 34

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
 1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
65 70 75 80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
85 90 95

Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His
100 105

<210> 35
<211> 223
<212> PRT
<213> Nhân tạo

<220>
<223> Cấu trúc tổng hợp

<400> 35

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
1 5 10 15

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
20 25 30

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Glu Pro Asp
35 40 45

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
50 55 60

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

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
85 90 95

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
100 105 110

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
115 120 125

Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
130 135 140

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
145 150 155 160

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
165 170 175

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
180 185 190

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
195 200 205

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
210 215 220

<210> 36
<211> 7814
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> immunoadhesin FI6 và CO5

<220>
<221> tín hiệu_polyA
<222> (201)..(432)
<223> bổ sung- SV40_polyA

<220>
<221> đặc điểm_đa dạng

<222> (453)..(1121)
 <223> bổ sung- CH'2-3

 <220>
 <221> đặc điểm đa dạng
 <222> (1125)..(1457)
 <223> bổ sung- C05\VL

 <220>
 <221> đặc điểm đa dạng
 <222> (1458)..(1502)
 <223> SL\from\3bn201co

 <220>
 <221> đặc điểm đa dạng
 <222> (1503)..(1916)
 <223> bổ sung- C05\VH

 <220>
 <221> đặc điểm đa dạng
 <222> (1965)..(1973)
 <223> đoạn dẫn

 <220>
 <221> đặc điểm đa dạng
 <222> (2371)..(2412)
 <223> bổ sung- CMV\mp2

 <220>
 <221> đặc điểm đa dạng
 <222> (2413)..(2881)
 <223> đoạn tăng cường

 <220>
 <221> đặc điểm đa dạng
 <222> (2882)..(3007)
 <223> CMV\mp1

 <220>
 <221> đặc điểm đa dạng
 <222> (3067)..(3055)
 <223> đoạn dẫn

 <220>
 <221> CDS
 <222> (3124)..(3510)
 <223> FI6\VH

 <220>
 <221> đặc điểm đa dạng
 <222> (3511)..(3555)
 <223> SL\from\3bn201co

 <220>

<221> CDS
 <222> (3556)..(3888)
 <223> FI6\VL

<220>
 <221> CDS
 <222> (3892)..(4560)
 <223> CH2-3

<220>
 <221> tín hiệu_polyA
 <222> (4581)..(4812)
 <223> SV40\polyA

<400> 36
 ggccttaatt aggctgcgcg ctgcgtcgct cactgaggcc gcccgggcaa agcccgggcg 60
 tcggggcgacc tttggctgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc 120
 caactccatc actaggggtt ccttgtagtt aatgattaac ccgccatgct acttatctac 180
 gtagccatgc tctaggaaga tcattttacc acattttagt aggttttact tgctttaaaa 240
 aacctccac atctcccct gaacctgaaa cataaaatga atgcaattgt tgttggttaac 300
 ttgtttattg cagcttataa tggttacaaa taaagcaata gcatcacaaa tttcacaaat 360
 aaagcatttt tttcactgca ttctagttgt ggtttgtcca aactcatcaa tgtatcttat 420
 catgtctgct cgaagcggcc gcaagcttat cacttcccgg ggctcaggct cagggacttc 480
 tgggtgtagt ggttgtgcag ggcctcgtgc atcacgtgc agctgaacac gttgccctgc 540
 tgccaccggc tcttgtccac ggtcagcttg ctatacagga agaattgagcc gtcgctgtcc 600
 agcacagggg ggggtggtctt gtagttgttc tcgggctggc cgttgctctc ccattccacg 660
 gcgatctcgc tggggtagaa gcccttgacc aggcaggta gggacacctg gttcttggtc 720
 atctcttccc ggctgggggg cagtgtgtag acctgaggct cgcggggctg gcccttggcc 780
 ttgctgatgg tttctctgat gggggcaggc agggccttgt tggacacctt gcacttgtag 840
 tctttgccgt tcagccagtc ctggtgcagc acggtcagca cggacaccac ccggtaggtg 900
 ctgttgtagt gttcctctct gggcttggtc ttggcgttgt gcacttccac gccgtccacg 960
 taccaattga acttcacttc agggctcctg tgggacacgt ccaccaccac gcaggtcact 1020
 tcgggggtcc ggctgatcat cagggtgtcc ttgggctttg gggggaacag gaacacgctg 1080
 gggcctccca gcagttcagg ggcagggcag gggggacacg tggctagcac cgtacgcttg 1140
 atctccagct tggttcctcc tccgaaggtg aatggcagtc catcgtactg ctggcagtag 1200

taggttccca catccttcag gctgcaggcc acgctgctgc tcaggctgat ctgtcccaga	1260
tccactccgc tgaaccggct tggcactccc cgctgcaggt tgctggcatc gtagatcagc	1320
agctttgggtc cctttcctgg cttctgctgg taccagttca ggaacttcct gatgtcctgg	1380
ctggcctggc aggtcagggt caccgatct cccacgctgg cgctcaggct gcttgggctc	1440
tgggtcagct ggatatcaga tccccgcct ccggaccctc ctccctccgct gcctcctccg	1500
ccgctcgaga cggtcaccag ggttccctgt cccacacat cgaaggcatc tcccaccaga	1560
tcggcccgtc cccatccggc gctcaccacc tgctgcatgg acatgtgctt ggcgcagtag	1620
tacactccgg tatectccac ccgcaggttg gtcattctga ggtacagggt ctcccttgctg	1680
ttatcccggc tgatggtgaa ccgtccctcc acgctatcgg cgtaatcaat gtctcctcct	1740
ccggcgttga tgatgctcag cactccagt ccctttcctg gggcctgccg caccagctc	1800
acggcgtagt agctcagggt gctctctccg aagctgcttc cgcttccac gcagctcagc	1860
cgcaggctct ctctggctg caccagtcct cctccgctct cctgcagctg cacctgtgaa	1920
ttcgtcacca gggccaggct caggcgcatc agcagcagca gctgcatgcg catggtgggg	1980
cggcagagtc gcgtccttgc tcgggtgttg taagttccag tgcaaagtgc cctagcctat	2040
agtgagtcgt attaagtact cttagcctta gagctgtaat tgaactggga gtggacacct	2100
gtggagagaa aggcaaagtg gatgtcagta agaccaatag gtgcctatca gaaacgcaag	2160
agtcttctct gtctcgacaa gccagtttc tattggtctc cttaaacctg tcttgtaacc	2220
ttgatactta cctgcccagt gcctcacgac caacttctgc agcttaagtt cgagactgtt	2280
gtgtcagaag cactgactgc gttagcaatt taactgtgat aaactaccgc aataaagctt	2340
ctagtgatct gacggttcac taaacgagct ctgcttatat aggctccca ccgtacacgc	2400
cacctcgaca tacctagtta ttaatagtaa tcaattacgg ggtcattagt tcatagccca	2460
tatatggagt tccgcgttac ataacttacg gtaaatggcc cgctggctg accgccaac	2520
gacccccgcc cattgacgtc aataatgacg tatgttccca tagtaacgcc aatagggact	2580
ttccattgac gtcaatgggt ggagtattta cggtaaactg ccacttggc agtacatcaa	2640
gtgtatcata tgccaagtac gccccctatt gacgtcaatg acggtaaagt gccgcctgg	2700
cattatgccc agtacatgac cttatgggac tttcctactt ggcagtacat ctacgtatta	2760
gtcatcgcta ttacatggt gatgcggttt tggcagtaca tcaatgggcg tggatagcgg	2820

tttgactcac ggggatttcc aagtctccac cccattgacg tcaatgggag tttgttttgg	2880
cacaaaaatc aacgggactt tccaaaatgt cgtaacaact ccgccccatt gacgcaaattg	2940
ggcggtaggc gtgtacggtg ggaggtctat ataagcagag ctggtttagt gaaccgtcag	3000
atccgctggg cactttgcac tggaacttac aacacccgag caaggacgcg actctgccgc	3060
cccaccatgc gcatgcagct gctgctgctg atcgccctga gcctggccct ggtgaccaac	3120
agc cag gtg caa ttg gtg gag agc gga gga gga gtg gtg cag cca gga Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly 1 5 10 15	3168
cgg agc ctg cgg ctg agc tgc gcc gcc agc gga ttc acc ttc agc acc Arg Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr 20 25 30	3216
tac gcc atg cac tgg gtg cgg cag gcc cca gga aag gga ctg gag tgg Tyr Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp 35 40 45	3264
gtg gcc gtg atc agc tac gat gcc aac tac aag tac tac gcc gat agc Val Ala Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser 50 55 60	3312
gtg aag gga cgg ttc acc atc agc cgg gat aac agc aag aac acc ctg Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu 65 70 75	3360
tac ctg cag atg aac agc ctg cgg gcc gag gat acc gcc gtg tac tac Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr 80 85 90 95	3408
tgc gcc aag gat agc cag ctg cgg agc ctg ctg tac ttc gag tgg ctg Cys Ala Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu 100 105 110	3456
agc cag gga tac ttc gat tac tgg gga cag gga acc ctg gtg acc gtg Ser Gln Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val 115 120 125	3504
agc agc ggcggaggag gcagcggagg aggaggggtcc ggaggcgggg gatct gat atc Ser Ser Asp Ile 130	3561
gtc atg acc cag agc cca gat agc ctg gcc gtg agc ctg gga gag cgg Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly Glu Arg 135 140 145	3609
gcc acc atc aac tgc aag agc agc cag agc gtg acc ttc aac tac aag Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Thr Phe Asn Tyr Lys 150 155 160	3657

aac tac ctg gcc tgg tac cag cag aag cca gga cag cca cca aag ctg	3705
Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu	
165 170 175	
ctg atc tac tgg gcc agc acc cgg gag agc gga gtg cca gat cgg ttc	3753
Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg Phe	
180 185 190 195	
agc gga agc gga agc gga acc gat ttc acc ctg acc atc agc agc ctg	3801
Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu	
200 205 210	
cag gcc gag gat gtg gcc gtg tac tac tgc cag cag cac tac cgg acc	3849
Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln His Tyr Arg Thr	
215 220 225	
cca cca acc ttc gga cag gga acc aag gtg gag atc aag gcc acg tgc	3897
Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Thr Cys	
230 235 240	
cct cca tgt cca gcc ccc gaa ctg ctg ggc ggg cct agc gtg ttc ctg	3945
Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu	
245 250 255	
ttt ccc cct aag cct aaa gat aca ctg atg att agt aga acc cca gag	3993
Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu	
260 265 270	
gtc aca tgc gtg gtc gtg gac gtg tcc cac gaa gag cct gac gtg aag	4041
Val Thr Cys Val Val Val Asp Val Ser His Glu Glu Pro Asp Val Lys	
275 280 285 290	
ttc aac tgg tac gtg gat ggc gtg gag gtg cac aat gct aag act aaa	4089
Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys	
295 300 305	
cca cgc gaa gag cag tat aat agt aca tac cga gtc gtg tca gtc ctg	4137
Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu	
310 315 320	
aca gtg ctg cac cag gat tgg ctg aac ggc aag gag tat aag tgc aag	4185
Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys	
325 330 335	
gtg tct aac aag gcc ctg ccc gcc cct atc gag aaa aca att agc aag	4233
Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys	
340 345 350	
gcc aaa ggg cag cca cgg gaa ccc cag gtc tac act ctg cca ccc tca	4281
Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser	
355 360 365 370	
aga gat gaa ctg act aag aac cag gtc agc ctg acc tgt ctg gtg aaa	4329

Arg	Asp	Glu	Leu	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys	Leu	Val	Lys	
				375					380					385		
ggc	ttc	tac	ccc	agc	gac	atc	gcc	gtg	gag	tgg	gaa	agt	aac	ggc	cag	4377
Gly	Phe	Tyr	Pro	Ser	Asp	Ile	Ala	Val	Glu	Trp	Glu	Ser	Asn	Gly	Gln	
			390					395					400			
cct	gag	aat	aac	tac	aag	act	acc	cct	cca	gtg	ctg	gat	agc	gac	ggg	4425
Pro	Glu	Asn	Asn	Tyr	Lys	Thr	Thr	Pro	Pro	Val	Leu	Asp	Ser	Asp	Gly	
			405				410					415				
tcc	ttc	ttc	ctg	tat	agc	aag	ctg	aca	gtg	gac	aaa	tcc	cgc	tgg	cag	4473
Ser	Phe	Phe	Leu	Tyr	Ser	Lys	Leu	Thr	Val	Asp	Lys	Ser	Arg	Trp	Gln	
	420					425				430						
cag	gga	aac	gtc	ttt	tcc	tgt	tct	gtg	atg	cat	gag	gcc	ctg	cac	aat	4521
Gln	Gly	Asn	Val	Phe	Ser	Cys	Ser	Val	Met	His	Glu	Ala	Leu	His	Asn	
435					440				445						450	
cat	tac	acc	cag	aag	agt	ctg	tca	ctg	agc	ccc	ggc	aaa	tgataagctt			4570
His	Tyr	Thr	Gln	Lys	Ser	Leu	Ser	Leu	Ser	Pro	Gly	Lys				
			455					460								
gcggccgctt	cgagcagaca	tgataagata	cattgatgag	tttggacaaa	ccacaactag											4630
aatgcagtga	aaaaaatgct	ttatattgtga	aatttgtgat	gctattgctt	tatttghtaac											4690
cattataagc	tgcaataaac	aagttaacaa	caacaattgc	attcatttta	tgttttcaggt											4750
tcagggggag	atgtggggagg	tttttttaaag	caagtaaaac	ctctacaaat	gtggtaaaat											4810
cgataaggat	cttcctagag	catggctacg	tagataagta	gcatggcggg	ttaatcatta											4870
actacaagga	acccttagtg	atggagtttg	ccactccctc	tctgcgcgct	cgctcgctca											4930
ctgaggccgg	gcgaccaaag	gtcgcccgac	gcccgggctt	tgcccgggcg	gcctcagtga											4990
gcgagcgagc	gcgcagcctt	aattaaccta	attcactggc	cgtcgtttta	caacgtcgtg											5050
actgggaaaa	ccctggcggt	acccaactta	atcgcccttg	agcacatccc	cctttcgcca											5110
gctggcgtaa	tagcgaagag	gcccgcaccg	atcgcccttc	ccaacagttg	cgcagcctga											5170
atggcgaaatg	ggacgcgccc	tgtagcggcg	cattaagcgc	ggcgggtgtg	gtggttacgc											5230
gcagcgtgac	cgctacactt	gccagcgcgc	tagcgcgcgc	tcctttcgct	ttcttccctt											5290
cctttctcgc	cacgttcgcc	ggctttcccc	gtcaagctct	aaatcggggg	ctcccttttag											5350
ggttccgatt	tagtgcttta	cggcacctcg	accccaaaaa	acttgattag	ggtgatgggt											5410
cacgtagtgg	gccatcgccc	tgatagacgg	tttttcgccc	tttgacgttg	gagtccacgt											5470
tctttaatatg	tggactcttg	ttccaaactg	gaacaacact	caaccctatc	tcgggtctatt											5530

cttttgattt ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt	5590
aacaaaaatt taacgcgaat ttttaacaaa tattaacgct tacaatttag gtggcacttt	5650
tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta	5710
tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat	5770
gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt	5830
ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg	5890
agtgggttac atcgaactgg atctcaacag cggtaaagtc cttgagagtt ttgcgccga	5950
agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcggt tattatcccg	6010
tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggt	6070
tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg	6130
cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg	6190
aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga	6250
tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc	6310
tgtagcaatg gcaacaacgt tgcgcaaact attactggc gaactactta ctctagcttc	6370
ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc	6430
ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg	6490
cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac	6550
gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc	6610
actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt	6670
aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac	6730
caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa	6790
aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc	6850
accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaagg	6910
aactggcttc agcagagcgc agatacaaaa tactgttctt ctagtgtagc cgtagttagg	6970
ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc	7030
agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt	7090
accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc ccagcttgga	7150

gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 7210
 tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg 7270
 cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 7330
 cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa 7390
 cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg ctacatggtt 7450
 ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga 7510
 taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga 7570
 gcgccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat gcagctggca 7630
 cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg tgagttagct 7690
 cactcattag gcaccccagg ctttacactt tatgcttccg gctcgtatgt tgtgtggaat 7750
 tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg ccagatttaa 7810
 ttaa 7814

<210> 37
 <211> 129
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 37

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
 1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ala Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser Val
 50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
 65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95

Ala Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu Ser
 100 105 110

Gln Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 115 120 125

Ser

<210> 38
 <211> 111
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 38

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Thr Phe Asn
 20 25 30

Tyr Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro
 35 40 45

Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp
 50 55 60

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser
 65 70 75 80

Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln His Tyr
 85 90 95

Arg Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
 100 105 110

<210> 39
 <211> 223
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 39

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
 1 5 10 15

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
 20 25 30

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Glu Pro Asp
 35 40 45

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
 50 55 60

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

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
 85 90 95

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
 100 105 110

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
 115 120 125

Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
 130 135 140

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
 145 150 155 160

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
 165 170 175

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
 180 185 190

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
 195 200 205

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 210 215 220

<210> 40
 <211> 7784
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> immunoadhesin FI6 và CR8033

<220>
 <221> tín hiệu_polyA
 <222> (201)..(432)
 <223> bổ sung- SV40\polyA

<220>
 <221> đặc điểm đa dạng
 <222> (453)..(1121)
 <223> bổ sung- CH'2-3

<220>
 <221> đặc điểm đa dạng
 <222> (1125)..(1460)
 <223> bổ sung- 033\VL

<220>
 <221> đặc điểm đa dạng
 <222> (1461)..(1505)
 <223> SL\from\3bn201co

<220>
 <221> đặc điểm đa dạng
 <222> (1506)..(1886)
 <223> bổ sung- 033\VH

<220>
 <221> đặc điểm đa dạng
 <222> (1935)..(1946)
 <223> bổ sung- đoạn dẫn

<220>

<221> đặc điểm đa dạng
 <222> (2341)..(2382)
 <223> bổ sung- CMV\mp2

<220>
 <221> đặc điểm đa dạng
 <222> (2383)..(2851)
 <223> đoạn tăng cường

<220>
 <221> đặc điểm đa dạng
 <222> (2852)..(2977)
 <223> CMV\mp1

<220>
 <221> đặc điểm đa dạng
 <222> (3073)..(3045)
 <223> đoạn dẫn

<220>
 <221> CDS
 <222> (3094)..(3480)
 <223> FI6\XH

<220>
 <221> đặc điểm đa dạng
 <222> (3481)..(3525)
 <223> SL\from\3bn201co

<220>
 <221> CDS
 <222> (3526)..(3858)
 <223> FI6\VL

<220>
 <221> đặc điểm đa dạng
 <222> (3862)..(4530)
 <223> CH2-3

<220>
 <221> tín hiệu polyA
 <222> (4551)..(4782)
 <223> SV40\polyA

<400> 40
 ggccttaatt aggctgcgcg ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg 60
 tcgggcgacc tttggctgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc 120
 caactccatc actagggggtt ccttgtagtt aatgattaac ccgcatgct acttatctac 180
 gtagccatgc tctaggaaga tcattttacc acattttagt aggttttact tgctttaaaa 240
 aacctccac atctccccct gaacctgaaa cataaaatga atgcaattgt tgttgttaac 300

ttgtttattg	cagcttataa	tggttacaaa	taaagcaata	gcatcacaaa	tttcacaaat	360
aaagcatttt	tttcaactgca	ttctagttgt	ggtttgtcca	aactcatcaa	tgtatcttat	420
catgtctgct	cgaagcggcc	gcaagcttat	cacttcccgg	ggctcaggct	cagggacttc	480
tgggtgtagt	ggttgtgcag	ggcctcgtgc	atcacgctgc	agctgaacac	gttgccctgc	540
tgccaccggc	tcttgtccac	ggtcagcttg	ctatacagga	agaatgagcc	gtcgctgtcc	600
agcacagggg	gggtggtctt	gtagttgttc	tcgggctggc	cgttgctctc	ccattccacg	660
gcgatctcgc	tggggtagaa	gcccttgacc	aggcaggtea	gggacacctg	gttcttggtc	720
atctcttccc	ggctgggggg	cagtgtgtag	acctgaggct	cgcggggctg	gcccttggcc	780
ttgctgatgg	ttttctcgat	gggggcaggc	agggccttgt	tggacacctt	gcacttgtac	840
tctttgccgt	tcagccagtc	ctggtgcagc	acggtcagca	cggacaccac	ccggtagggtg	900
ctgttgtagt	gttccctctc	gggcttggtc	ttggcgttgt	gcacttccac	gccgtccacg	960
taccaattga	acttcacttc	agggtccctc	tgggacacgt	ccaccaccac	gcaggtcact	1020
tcgggggtcc	ggctgatcat	cagggtgtcc	ttgggccttg	gggggaacag	gaacacgctg	1080
gggcctccca	gcagttcagg	ggcagggcag	gggggacacg	tggctagcac	cgtacgcttg	1140
atctccacct	tggttccctg	tccgaaggtc	caagggtgc	ttccgtactg	ctggcagtag	1200
tacacggcca	gacccctctg	ctccagccgg	ctgatggtca	gggtgaaatc	ggttccgctt	1260
ccgcttccgc	tgaaccgggc	tgggattccg	gtggcccggg	tgctggctcc	gtagatcagc	1320
agccgtgggg	cctgtccctg	cttctgctgg	taccaggcca	ggtagctgct	gctcacgctc	1380
tggctggccc	ggcagctcag	ggtggcccgc	tctcctgggc	tcaggctcag	ggttccctggg	1440
ctctgggtca	gcacgatctc	agatcccccg	cctccggacc	ctcctcctcc	gctgcctcct	1500
ccgccgctgc	tcacggtcac	catggttccc	tgtccccaga	tatcgaaggt	tcctccctcc	1560
aggatatcca	tggcgctgct	ctccagccga	tccttggcgc	agtagtacag	ggcggtatcc	1620
tcggcccgc	ggctgttcat	ctgcaggtag	aggctgttct	ttccgttatc	ccggctgatg	1680
gtgaaccgct	cctgcacgct	atcggcgtag	cccatgaagt	ttcccttcca	gttgattccg	1740
gccaccact	ccagtccctt	tcctggggcc	tgccgcaccc	agtgcaggt	gtactcatcg	1800
aagctgaatc	cgtgggcggc	gcagctcagc	cgcaggctcc	gtcctggctg	caccagtcct	1860
cctccggtct	ccaccagctg	cacctctgaa	ttcgtcacca	gggccaggct	cagggcgatc	1920

```

agcagcagca gctgcatgcg catggtgggg cggcagagtc gcgtccttgc tcgggtgttg      1980
taagttccag tgcaaagtgc cctagcctat agtgagtcgt attaagtact ctagccttaa      2040
gagctgtaat tgaactggga gtggacacct gtggagagaa aggcaaagtg gatgtcagta      2100
agaccaatag gtgcctatca gaaacgcaag agtcttctct gtctcgacaa gccagtttc      2160
tattggtctc cttaaacctg tcttgtaacc ttgatactta cctgcccagt gcctcacgac      2220
caacttctgc agcttaagtt cgagactggt gtgtcagaag cactgactgc gttagcaatt      2280
taactgtgat aaactaccgc aataaagctt ctagtgatct gacggttcac taaacgagct      2340
ctgcttatat aggctccca ccgtacacgc cacctcgaca tacctagtta ttaatagtaa      2400
tcaattacgg ggtcattagt tcatagccca tatatggagt tccgcgttac ataacttacg      2460
gtaaattggcc cgcttggtcg accgccaac gacccccgcc cattgacgtc aataatgacg      2520
tatgttccca tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta      2580
cggtaaactg ccacttggc agtacatcaa gtgtatcata tgccaagtac gcccctatt      2640
gacgtcaatg acggtaaagt gcccgcctgg cattatgcc agtacatgac cttatgggac      2700
tttctactt ggcagtacat ctacgtatta gtcacgcta ttaccatggt gatgcggttt      2760
tggcagtaca tcaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac      2820
cccattgacg tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt      2880
cgtaacaact ccgccccatt gacgcaaagt ggcggtaggc gtgtacggtg ggaggctctat      2940
ataagcagag ctggtttagt gaaccgtcag atccgctggg cactttgcac tggaacttac      3000
aacacccgag caaggacgcg actctgccgc ccaccatgc gcatgcagct gctgctgctg      3060
atcgccctga gcctggccct ggtgaccaac agc cag gtg caa ttg gtg gag agc      3114
                               Gln Val Gln Leu Val Glu Ser
                               1               5

gga gga gga gtg gtg cag cca gga cgg agc ctg cgg ctg agc tgc gcc      3162
Gly Gly Gly Val Val Gln Pro Gly Arg Ser Leu Arg Leu Ser Cys Ala
          10                15                20

gcc agc gga ttc acc ttc agc acc tac gcc atg cac tgg gtg cgg cag      3210
Ala Ser Gly Phe Thr Phe Ser Thr Tyr Ala Met His Trp Val Arg Gln
          25                30                35

gcc cca gga aag gga ctg gag tgg gtg gcc gtg atc agc tac gat gcc      3258
Ala Pro Gly Lys Gly Leu Glu Trp Val Ala Val Ile Ser Tyr Asp Ala
          40                45                50                55

```

aac tac aag tac tac gcc gat agc gtg aag gga cgg ttc acc atc agc Asn Tyr Lys Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser 60 65 70	3306
cgg gat aac agc aag aac acc ctg tac ctg cag atg aac agc ctg cgg Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg 75 80 85	3354
gcc gag gat acc gcc gtg tac tac tgc gcc aag gat agc cag ctg cgg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Lys Asp Ser Gln Leu Arg 90 95 100	3402
agc ctg ctg tac ttc gag tgg ctg agc cag gga tac ttc gat tac tgg Ser Leu Leu Tyr Phe Glu Trp Leu Ser Gln Gly Tyr Phe Asp Tyr Trp 105 110 115	3450
gga cag gga acc ctg gtg acc gtg agc agc ggcggaggag gcagcggagg Gly Gln Gly Thr Leu Val Thr Val Ser Ser 120 125	3500
aggaggggtcc ggaggcgggg gatct gat atc gtc atg acc cag agc cca gat Asp Ile Val Met Thr Gln Ser Pro Asp 130 135	3552
agc ctg gcc gtg agc ctg gga gag cgg gcc acc atc aac tgc aag agc Ser Leu Ala Val Ser Leu Gly Glu Arg Ala Thr Ile Asn Cys Lys Ser 140 145 150	3600
agc cag agc gtg acc ttc aac tac aag aac tac ctg gcc tgg tac cag Ser Gln Ser Val Thr Phe Asn Tyr Lys Asn Tyr Leu Ala Trp Tyr Gln 155 160 165 170	3648
cag aag cca gga cag cca cca aag ctg ctg atc tac tgg gcc agc acc Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr 175 180 185	3696
cgg gag agc gga gtg cca gat cgg ttc agc gga agc gga agc gga acc Arg Glu Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr 190 195 200	3744
gat ttc acc ctg acc atc agc agc ctg cag gcc gag gat gtg gcc gtg Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val 205 210 215	3792
tac tac tgc cag cag cac tac cgg acc cca cca acc ttc gga cag gga Tyr Tyr Cys Gln Gln His Tyr Arg Thr Pro Pro Thr Phe Gly Gln Gly 220 225 230	3840
acc aag gtg gag atc aag gccacgtgcc ctccatgtcc agccccgaa Thr Lys Val Glu Ile Lys 235 240	3888
ctgctgggcg ggccatagcgt gttcctgttt cccctaagc ctaaagatac actgatgatt	3948

agtagaacc	cagaggtcac	atgcgtggtc	gtggacgtgt	cccacgaaga	gcctgacgtg	4008
aagttcaact	ggtacgtgga	tggcgtggag	gtgcacaatg	ctaagactaa	accacgcgaa	4068
gagcagtata	atagtacata	ccgagtcgtg	tcagtcctga	cagtgctgca	ccaggattgg	4128
ctgaacggca	aggagtataa	gtgcaagggtg	tctaacaagg	ccctgcccgc	ccctatcgag	4188
aaaacaatta	gcaaggccaa	agggcagcca	cgggaacccc	aggtctacac	tctgccaccc	4248
tcaagagatg	aactgactaa	gaaccagggtc	agcctgacct	gtctggtgaa	aggctttctac	4308
cccagcgaca	tcgccgtgga	gtgggaaagt	aacggccagc	ctgagaataa	ctacaagact	4368
accctccag	tgctggatag	cgacgggtcc	ttcttcctgt	atagcaagct	gacagtggac	4428
aaatcccgtc	ggcagcaggg	aaacgtcttt	tcctgttctg	tgatgcatga	ggccctgcac	4488
aatcattaca	cccagaagag	tctgtcactg	agccccggca	aatgataagc	ttgcggccgc	4548
ttcgagcaga	catgataaga	tacattgatg	agtttgga	aaccacaact	agaatgcagt	4608
gaaaaaatg	ctttatattgt	gaaatttgtg	atgctattgc	tttatttgta	accattataa	4668
gctgcaataa	acaagttaac	aacaacaatt	gcattcattt	tatgtttcag	gttcaggggg	4728
agatgtggga	ggttttttaa	agcaagtaaa	acctctacaa	atgtggtaaa	atcgataagg	4788
atcttcctag	agcatggcta	cgtagataag	tagcatggcg	ggttaatcat	taactacaag	4848
gaacccttag	tgatggagtt	ggccactccc	tctctgcgcg	ctcgctcgct	cactgaggcc	4908
gggcgaccaa	aggtcgcccg	acgcccgggc	tttgcccggg	cggcctcagt	gagcgagcga	4968
gcgcgcagcc	ttaattaacc	taattcactg	gccgtcgttt	tacaacgtcg	tgactgggaa	5028
aaccctggcg	ttaccacaact	taatgcctt	gcagcacatc	cccctttcgc	cagctggcgt	5088
aatagcgaag	aggcccgcac	cgatcgccct	tccaacagt	tgcgcagcct	gaatggcgaa	5148
tgggacgcgc	cctgtagcgg	cgcattaagc	gcggcgggtg	tggtggttac	gcgcagcgtg	5208
accgctacac	ttgccagcgc	cctagcgccc	gtccttttcg	ctttcttccc	ttcctttctc	5268
gccacgttcg	ccggctttcc	ccgtcaagct	ctaaatcggg	ggctcccttt	agggttccga	5328
tttagtgctt	tacggcacct	cgaccccaaa	aaacttgatt	agggtgatgg	ttcacgtagt	5388
gggccatcgc	cctgatagac	ggtttttcgc	cctttgacgt	tggagtccac	gttctttaat	5448
agtggaactct	tgttccaaac	tggaacaaca	ctcaacccta	tctcgggtcta	ttcttttgat	5508
ttataaggga	ttttgccgat	ttcggcctat	tggttaaaaa	atgagctgat	ttaacaaaaa	5568

tttaacgcga	attttaacaa	aatattaacg	cttacaat	aggtggcact	tttcggggaa	5628
atgtgcgcgg	aacccttatt	tgtttat	tctaaataca	ttcaaata	tatccgctca	5688
tgagacaata	accctgataa	atgcttcaat	aatattgaaa	aaggaagagt	atgagtattc	5748
aacatttccg	tgtcgccctt	attccctttt	ttgcggcatt	ttgccttcct	gtttttgctc	5808
accagaaac	gctggtgaaa	gtaaaagatg	ctgaagatca	gttggtgca	cgagtgggtt	5868
acatcgaact	ggatctcaac	agcggtaaga	tccttgagag	ttttcgcccc	gaagaacgtt	5928
ttccaatgat	gagcactttt	aaagttctgc	tatgtggcgc	ggtattatcc	cgtattgacg	5988
ccgggcaaga	gcaactcggt	cgccgcatac	actattctca	gaatgacttg	gttgagtact	6048
caccagtcac	agaaaagcat	cttacggatg	gcatgacagt	aagagaatta	tgcagtgctg	6108
ccataaccat	gagtataac	actgcggcca	acttacttct	gacaacgatc	ggaggaccga	6168
aggagctaac	cgcttttttg	cacaacatgg	gggatcatgt	aactcgcctt	gatcgttggg	6228
aaccggagct	gaatgaagcc	ataccaaacg	acgagcgtga	caccacgatg	cctgtagcaa	6288
tggcaacaac	gttgcgcaaa	ctattaactg	gcgaactact	tactctagct	tcccggaac	6348
aattaataga	ctggatggag	gcggataaag	ttgcaggacc	acttctgcgc	tcggcccttc	6408
cggctggctg	gtttattgct	gataaatctg	gagccggtga	gcgtgggtct	cgcggtatca	6468
ttgcagcact	ggggccagat	ggtaagccct	cccgtatcgt	agttatctac	acgacgggga	6528
gtcaggcaac	tatggatgaa	cgaaatagac	agatcgctga	gatagggtgcc	tactgatta	6588
agcattggta	actgtcagac	caagtttact	catatatact	ttagattgat	ttaaaacttc	6648
atttttaatt	taaaaggatc	taggtgaaga	tcctttttga	taatctcatg	acaaaaatcc	6708
cttaacgtga	gttttcgttc	cactgagcgt	cagaccccg	agaaaagatc	aaaggatc	6768
cttgagatcc	ttttttctg	cgcgtaatct	gctgcttgca	aacaaaaaaa	ccaccgctac	6828
cagcgggtgg	ttgtttgccg	gatcaagagc	taccaactct	ttttccgaag	gtaactggct	6888
tcagcagagc	gcagatacca	aatactgttc	ttctagtgtg	gccgtagtta	ggccaccact	6948
tcaagaactc	tgtagcaccg	cctacatacc	tcgctctgct	aatcctgtta	ccagtggctg	7008
ctgccagtgg	cgataagtcg	tgtcttaccg	ggttgactc	aagacgatag	ttaccggata	7068
aggcgcagcg	gtcgggctga	acgggggggt	cgtgcacaca	gccagcttg	gagcgaacga	7128
cctacaccga	actgagatac	ctacagcgtg	agctatgaga	aagcgccacg	cttcccgaag	7188

ggagaaaggc ggacaggtat ccggttaagcg gcagggtcgg aacaggagag cgcacgaggg 7248
 agcttccagg gggaaacgcc tggatatcttt atagtcctgt cgggtttcgc cacctctgac 7308
 ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca 7368
 acgcggcctt tttacgggtc ctggcctttt gctggccttt tgctcacatg ttctttcctg 7428
 cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgaagt gataccgctc 7488
 gccgcagccg aacgaccgag cgcagcgagt cagtgaagca ggaagcggaa gagcgcccaa 7548
 tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg cacgacaggt 7608
 ttcccgaactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag ctactcatt 7668
 aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga attgtgagcg 7728
 gataacaatt tcacacagga aacagctatg accatgatta cgccagattt aattaa 7784

<210> 41
 <211> 129
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 41

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
 1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ala Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser Val
 50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
 65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95

Ala Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu Ser
 100 105 110

Gln Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 115 120 125

Ser

<210> 42
 <211> 111
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 42

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Thr Phe Asn
 20 25 30

Tyr Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro
 35 40 45

Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp
 50 55 60

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser
 65 70 75 80

Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln His Tyr
 85 90 95

Arg Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
 100 105 110

<210> 43
 <211> 7782

<212> ADN
 <213> Trình tự nhân tạo

 <220>
 <223> Plasmit mang kháng thể đơn dòng TCN032 và Fi6

 <220>
 <221> vùng_lặp
 <222> (14)..(143)

 <220>
 <221> tín hiệu_polyA
 <222> (204)..(252)
 <223> tổng hợp polyA

 <220>
 <221> đặc điểm_đa dạng
 <222> (261)..(267)
 <223> cat-xet chặn (bổ sung)

 <220>
 <221> đặc điểm_đa dạng
 <222> (268)..(588)
 <223> chuỗi nhẹ hằng định (trên chuỗi bổ sung)

 <220>
 <221> đặc điểm_đa dạng
 <222> (967)..(971)
 <223> Kozak (nằm trên chuỗi bổ sung)

 <220>
 <221> đặc điểm_đa dạng
 <222> (972)..(1019)
 <223> c-myc 5' UTR (nằm trên chuỗi bổ sung)

 <220>
 <221> đặc điểm_đa dạng
 <222> (1026)..(1094)
 <223> CMV\mp2

 <220>
 <221> đoạn tăng cường
 <222> (1026)..(1094)

 <220>
 <221> đặc điểm_đa dạng
 <222> (1564)..(1689)

 <220>
 <221> đặc điểm_đa dạng
 <222> (1696)..(1743)
 <223> c-myc 5' UTR

<220>
 <221> đặc điểm đa dạng
 <222> (1744)..(1748)
 <223> Kozak

<220>
 <221> đặc điểm đa dạng
 <222> (1749)..(1805)
 <223> đoạn dẫn

<220>
 <221> CDS
 <222> (1806)..(2165)
 <223> Chuỗi nặng biến đổi TCN032

<220>
 <221> vùng_lặp
 <222> (1845)..(4974)
 <223> đoạn lặp tận cùng đảo ngược

<220>
 <221> vùng_lặp
 <222> (1845)..(4974)
 <223> đoạn lặp tận cùng đảo ngược (nằm trên chuỗi bổ sung)

<220>
 <221> CDS
 <222> (2166)..(2459)
 <223> CH1

<220>
 <221> misc
 <222> (2166)..(2459)
 <223> CH1

<220>
 <221> CDS
 <222> (2460)..(3152)
 <223> khớp nối-CH2'-CH3'

<220>
 <221> đặc điểm đa dạng
 <222> (3153)..(3164)
 <223> vị trí phân cắt furin

<220>
 <221> đặc điểm đa dạng
 <222> (3165)..(3236)
 <223> Đoạn liên kết F2A

<220>
 <221> đặc điểm đa dạng
 <222> (3239)..(3296)

<220>
 <221> đặc điểm đa dạng
 <222> (3239)..(3296)

<220>
 <221> CDS
 <222> (3297)..(3683)
 <223> FI6 VH

<220>
 <221> CDS
 <222> (3684)..(4004)
 <223> CH1

<220>
 <221> CDS
 <222> (4005)..(4673)
 <223> CH2-3

<220>
 <221> đặc điểm đa dạng
 <222> (4674)..(4680)
 <223> Cat-xet chặn

<220>
 <221> đặc điểm đa dạng
 <222> (4674)..(4680)

<220>
 <221> tín hiệu polyA
 <222> (4693)..(4770)
 <223> TKpA ngắn

<220>
 <221> rep_gốc
 <222> (5151)..(5606)

<220>
 <221> CDS
 <222> (5737)..(6594)
 <223> Amp-R

<220>
 <221> đặc điểm đa dạng
 <222> (6768)..(7356)
 <223> col\E1\gốc

<400> 43
 ggccttaatt aggctgcgcg ctcgctcgct cactgaggcc gcccgggcaa agcccgggcg 60
 tcgggcgacc tttggctgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc 120
 caactccatc actagggggt ccttgtagtt aatgattaac ccgcatgct acttatctac 180

gtagccatgc tctaggaaga tctcacacaa aaaaccaaca cacagatgta atgaaaataa	240
agatatTTTta ttgoggccgc tttatcagca ctctccgcgg ttgaagctct tggtcactgg	300
gctgctcagt ccctggtggg tcacctcgca ggcgtacacc ttgtgcttct cgtaatcggc	360
cttgctcagg gtcaggggtgc tgctcaggct gtaggtgcta tccttgctat cctgctcggc	420
cacgctctcc tggctgtttc cgctctgcag ggcgttatcc accttcact gcaccttggc	480
ctcccgtggg tagaagttgt tcagcaggca caccacgctg gcggttccgc tcttcagctg	540
ctcatcgctt ggtgggaaga tgaacacgct tggggcgcc accgtacgct tgatctccac	600
ccgggttcct cctccgaagg tcagtgggtg gctgtagctc tgctggcagt agtaggtggc	660
gaaatcctct ggctgcaggc tggatgatgg cagggtgaaa tcggttccgc ttccgcttcc	720
gctgaaccgg cttggcactc cgctctgcag tccgctggcg gcgctgatca gtccctttgg	780
ggcctttcct ggccgctgct ggtaccagtt caggtacttg tagatgttct ggctggcccg	840
gcaggtgatg gtcacccgat ctcccacgct ggcgctcagg ctgcttgggc tctgggtcat	900
ctggatatcg ctggttgtca ccagggccag gctcagggcg atcagcagca gcagctgcat	960
tctcatggtg gagagtcgcg tccttgctcg ggtgttgtaa gttccagtgc aaagtgcccc	1020
aattggcgat ctgacggttc actaaacgag ctctgcttat ataggcctcc caccgtacac	1080
gccacctcga catacctagt tattaatagt aatcaattac ggggtcatta gttcatagcc	1140
catatatgga gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgcca	1200
acgacccccg ccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga	1260
ctttccattg acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc	1320
aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgct	1380
ggcattatgc ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat	1440
tagtcatcgc tattaccatg gtgatgcggg tttggcagta catcaatggg cgtggatagc	1500
ggtttgactc acggggatTTT ccaagtctcc accccattga cgtcaatggg agtttgtttt	1560
ggcaccaaaa tcaacgggac tttccaaaat gtcgtaacaa ctccgcccc ttgacgcaaa	1620
tgggcggtag gcgtgtacgg tgggaggtct atataagcag agctggttta gtgaaccgtc	1680
agatccgctg ctagcgggca ctttgactg gaacttaca caccgagca aggacgcgac	1740
tctccaccat gcgcatgcag ctgctgctgc tgatcgccct gagcctggcc ctggtgacca	1800

acagc	cag	gtg	cag	ctg	cag	gag	agc	gga	cca	gga	ctg	gtg	aag	cca	agc	1850
	Gln	Val	Gln	Leu	Gln	Glu	Ser	Gly	Pro	Gly	Leu	Val	Lys	Pro	Ser	
	1				5					10					15	
gag	acc	ctg	agc	ctg	acc	tgc	acc	gtg	agc	gga	agc	agc	atc	agc	aac	1898
Glu	Thr	Leu	Ser	Leu	Thr	Cys	Thr	Val	Ser	Gly	Ser	Ser	Ile	Ser	Asn	
				20					25					30		
tac	tac	tgg	agc	tgg	atc	cgg	cag	agc	cca	gga	aag	gga	ctg	gag	tgg	1946
Tyr	Tyr	Trp	Ser	Trp	Ile	Arg	Gln	Ser	Pro	Gly	Lys	Gly	Leu	Glu	Trp	
			35					40					45			
atc	gga	ttc	atc	tac	tac	gga	gga	aac	acc	aag	tac	aac	cca	agc	ctg	1994
Ile	Gly	Phe	Ile	Tyr	Tyr	Gly	Gly	Asn	Thr	Lys	Tyr	Asn	Pro	Ser	Leu	
		50					55					60				
aag	agc	cgg	gtg	acc	atc	agc	cag	gat	acc	agc	aag	agc	cag	gtg	agc	2042
Lys	Ser	Arg	Val	Thr	Ile	Ser	Gln	Asp	Thr	Ser	Lys	Ser	Gln	Val	Ser	
	65					70					75					
ctg	acc	atg	agc	agc	gtg	acc	gcc	gcc	gag	agc	gcc	gtg	tac	ttc	tgc	2090
Leu	Thr	Met	Ser	Ser	Val	Thr	Ala	Ala	Glu	Ser	Ala	Val	Tyr	Phe	Cys	
80					85					90					95	
gcc	cgg	gcc	agc	tgc	agc	gga	gga	tac	tgc	atc	ctg	gat	tac	tgg	gga	2138
Ala	Arg	Ala	Ser	Cys	Ser	Gly	Gly	Tyr	Cys	Ile	Leu	Asp	Tyr	Trp	Gly	
				100					105					110		
cag	gga	acc	ctg	gtg	acc	gtg	agc	agc	gcg	tcg	acc	aag	gga	cct	tcg	2186
Gln	Gly	Thr	Leu	Val	Thr	Val	Ser	Ser	Ala	Ser	Thr	Lys	Gly	Pro	Ser	
			115					120					125			
gtc	ttc	ccc	ctg	gca	ccc	tcc	tcc	aag	agc	acc	tct	ggg	ggc	aca	gcg	2234
Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys	Ser	Thr	Ser	Gly	Gly	Thr	Ala	
		130					135					140				
gcc	ctg	ggc	tgc	ctg	gtc	aag	gac	tac	ttc	ccc	gaa	ccg	gtg	acg	gtg	2282
Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	Phe	Pro	Glu	Pro	Val	Thr	Val	
	145					150					155					
tcg	tgg	aac	tca	ggc	gcc	ctg	acc	agc	ggc	gtg	cac	acc	ttc	ccg	gct	2330
Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	Gly	Val	His	Thr	Phe	Pro	Ala	
160					165					170					175	
gtc	cta	cag	tcc	tca	gga	ctc	tac	tcc	ctc	agc	agc	gtg	gtg	acc	gtg	2378
Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val	Thr	Val	
				180					185					190		
ccc	tcc	agc	agc	ttg	ggc	acc	cag	acc	tac	atc	tgc	aac	gtg	aat	cac	2426
Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	Cys	Asn	Val	Asn	His	
				195				200					205			
aag	ccc	agc	aac	acc	aag	gtg	gac	aag	aaa	gtt	gaa	cca	aag	agc	tgc	2474
Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	Glu	Pro	Lys	Ser	Cys	

210	215	220	
gac aag acc cac acg tgt ccc ccc tgc cct gcc cct gaa ctg ctg gga Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly 225 230 235			2522
ggc ccc agc gtg ttc ctg ttc ccc cca aag ccc aag gac acc ctg atg Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met 240 245 250 255			2570
atc agc cgg acc ccc gaa gtg acc tgc gtg gtg gtg gac gtg tcc cac Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His 260 265 270			2618
gag gac cct gaa gtg aag ttt aat tgg tac gtg gac ggc gtg gaa gtg Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val 275 280 285			2666
cac aac gcc aag acc aag ccc aga gag gaa cag tac aac agc acc tac His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr 290 295 300			2714
cgg gtg gtg tcc gtg ctg acc gtg ctg cac cag gac tgg ctg aac ggc Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly 305 310 315			2762
aaa gag tac aag tgc aag gtg tcc aac aag gcc ctg cct gcc ccc atc Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile 320 325 330 335			2810
gag aaa acc atc agc aag gcc aag ggc cag ccc cgc gag cct cag gtc Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val 340 345 350			2858
tac aca ctg ccc ccc agc cgg gaa gag atg acc aag aac cag gtg tcc Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser 355 360 365			2906
ctg acc tgc ctg gtc aag ggc ttc tac ccc agc gac atc gcc gtg gaa Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu 370 375 380			2954
tgg gag agc aac ggc cag ccc gag aac aac tac aag acc acc ccc cct Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro 385 390 395			3002
gtg ctg gac agc gac ggc tca ttc ttc ctg tat agc aag ctg acc gtg Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val 400 405 410 415			3050
gac aag agc cgg tgg cag cag ggc aac gtg ttc agc tgc agc gtg atg Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met 420 425 430			3098

cac gag gcc ctg cac aac cac tac acc cag aag tcc ctg agc ctg agc 3146
 His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

ccc ggc agaaagcgga gagccccctg gaagcagacc ctgaacttcg acctgctgaa 3202
 Pro Gly

gctggccggc gacgtggaaa gcaaccctgg ccctatgtac agaatgcagc tgctgagctg 3262

catcgccctg agcctggccc tggtagaccaa cagc cag gtg caa cta gtg gag agc 3317
 Gln Val Gln Leu Val Glu Ser
 450 455

gga gga gga gtg gtg cag cca gga cgg agc ctg cgg ctg agc tgc gcc 3365
 Gly Gly Gly Val Val Gln Pro Gly Arg Ser Leu Arg Leu Ser Cys Ala
 460 465 470

gcc agc gga ttc acc ttc agc acc tac gcc atg cac tgg gtg cgg cag 3413
 Ala Ser Gly Phe Thr Phe Ser Thr Tyr Ala Met His Trp Val Arg Gln
 475 480 485

gcc cca gga aag gga ctg gag tgg gtg gcc gtg atc agc tac gat gcc 3461
 Ala Pro Gly Lys Gly Leu Glu Trp Val Ala Val Ile Ser Tyr Asp Ala
 490 495 500

aac tac aag tac tac gcc gat agc gtg aag gga cgg ttc acc atc agc 3509
 Asn Tyr Lys Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser
 505 510 515 520

cgg gat aac agc aag aac acc ctg tac ctg cag atg aac agc ctg cgg 3557
 Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg
 525 530 535

gcc gag gat acc gcc gtg tac tac tgc gcc aag gat agc cag ctg cgg 3605
 Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Lys Asp Ser Gln Leu Arg
 540 545 550

agc ctg ctg tac ttc gag tgg ctg agc cag gga tac ttc gat tac tgg 3653
 Ser Leu Leu Tyr Phe Glu Trp Leu Ser Gln Gly Tyr Phe Asp Tyr Trp
 555 560 565

gga cag gga acc ctg gtg acc gtg agc agc gcc agc acc aag ggg ccc 3701
 Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro
 570 575 580

agc gtg ttc cca ctg gcc cca agc agc aag agc acc agc gga gga acc 3749
 Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr
 585 590 595 600

gcc gcc ctg gga tgc ctg gtg aag gat tac ttc cca gag cca gtg acc 3797
 Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr
 605 610 615

gtg agc tgg aac agc gga gcc ctg acc agc gga gtg cac acc ttc cca Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro 620 625 630	3845
gcc gtg ctg cag agc agc gga ctg tat agc ctg agc agc gtg gtg acc Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr 635 640 645	3893
gtg cca agc agc agc ctg gga acc cag acc tac atc tgc aac gtg aac Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn 650 655 660	3941
cac aag cca agc aac acc aag gtg gat aag aag gtg gag cca aag agc His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser 665 670 675 680	3989
tgc gat aag acc cac acg tgc cct cca tgt cca gcc ccc gaa ctg ctg Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu 685 690 695	4037
ggc ggg cct agc gtg ttc ctg ttt ccc cct aag cct aaa gat aca ctg Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu 700 705 710	4085
atg att agt aga acc cca gag gtc aca tgc gtg gtc gtg gac gtg tcc Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser 715 720 725	4133
cac gaa gag cct gac gtg aag ttc aac tgg tac gtg gat ggc gtg gag His Glu Glu Pro Asp Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu 730 735 740	4181
gtg cac aat gct aag act aaa cca cgc gaa gag cag tat aat agt aca Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr 745 750 755 760	4229
tac cga gtc gtg tca gtc ctg aca gtg ctg cac cag gat tgg ctg aac Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn 765 770 775	4277
ggc aag gag tat aag tgc aag gtg tct aac aag gcc ctg ccc gcc cct Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro 780 785 790	4325
atc gag aaa aca att agc aag gcc aaa ggg cag cca cgg gaa ccc cag Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln 795 800 805	4373
gtc tac act ctg cca ccc tca aga gat gaa ctg act aag aac cag gtc Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val 810 815 820	4421
agc ctg acc tgt ctg gtg aaa ggc ttc tac ccc agc gac atc gcc gtg Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val	4469

825	830	835	840	
gag tgg gaa agt aac ggc cag cct gag aat aac tac aag act acc cct				4517
Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro				
	845	850	855	
cca gtg ctg gat agc gac ggg tcc ttc ttc ctg tat agc aag ctg aca				4565
Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr				
	860	865	870	
gtg gac aaa tcc cgc tgg cag cag gga aac gtc ttt tcc tgt tct gtg				4613
Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val				
	875	880	885	
atg cat gag gcc ctg cac aat cat tac acc cag aag agt ctg tca ctg				4661
Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu				
	890	895	900	
agc ccc ggc aaa tgataaaaag cttctcgaga aggaacccgc gctatgacgg				4713
Ser Pro Gly Lys				
905				
caataaaaag acagaataaa acccacgggt gttgggtcgt ttgttcataa acccggaag				4773
cttatcgata aggatcttcc tagagcatgg ctacgtagat aagtagcatg gcgggttaat				4833
cattaactac aaggaacccc tagtgatgga gttggccact ccctctctgc gcgctcgctc				4893
gctcactgag gccgggcgac caaaggctgc cgcacgcccg ggctttgccc gggcggcctc				4953
agtgagcgag cgagcgcgca gccttaatta acctaattca ctggccgtcg ttttacaacg				5013
tcgtgactgg gaaaaccctg gcgttaccca acttaatcgc cttgcagcac atcccccttt				5073
cgccagctgg cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag				5133
cctgaatggc gaatgggacg cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt				5193
tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt				5253
cccttccttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc				5313
tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg attagggta				5373
tggttcacgt agtgggccat cgccctgata gacggttttt cgccctttga cgttggagtc				5433
cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc ctatctcgg				5493
ctattctttt gatattataag ggattttgcc gatttcggcc tattgggttaa aaaatgagct				5553
gatttaacaa aaatttaacg cgaattttta caaaatatta acgcttacia tttaggtggc				5613
acttttcggg gaaatgtgcg cggaaccctt atttgtttat ttttctaaat acattcaaat				5673

atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag	5733
agt atg agt att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala 910 915 920	5781
gca ttt tgc ctt cct gtt ttt gct cac cca gaa acg ctg gtg aaa gta Ala Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val 925 930 935	5829
aaa gat gct gaa gat cag ttg ggt gca cga gtg ggt tac atc gaa ctg Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu 940 945 950 955	5877
gat ctc aac agc ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg 960 965 970	5925
ttt cca atg atg agc act ttt aaa gtt ctg cta tgt ggc gcg gta tta Phe Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu 975 980 985	5973
tcc cgt att gac gcc ggg caa gag caa ctc ggt cgc cgc ata cac tat Ser Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr 990 995 1000	6021
tct cag aat gac ttg gtt gag tac tca cca gtc aca gaa aag cat Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His 1005 1010 1015	6066
ctt acg gat ggc atg aca gta aga gaa tta tgc agt gct gcc ata Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile 1020 1025 1030	6111
acc atg agt gat aac act gcg gcc aac tta ctt ctg aca acg atc Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile 1035 1040 1045	6156
gga gga ccg aag gag cta acc gct ttt ttg cac aac atg ggg gat Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp 1050 1055 1060	6201
cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa gcc His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala 1065 1070 1075	6246
ata cca aac gac gag cgt gac acc acg atg cct gta gca atg gca Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val Ala Met Ala 1080 1085 1090	6291
aca acg ttg cgc aaa cta tta act ggc gaa cta ctt act cta gct Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala 1095 1100 1105	6336

tcc cgg	caa caa tta ata gac	tgg atg gag gcg gat	aaa gtt gca	6381		
Ser Arg	Gln Gln Leu Ile Asp	Trp Met Glu Ala Asp	Lys Val Ala			
1110	1115	1120				
gga cca	ctt ctg cgc tcg gcc	ctt ccg gct ggc tgg	ttt att gct	6426		
Gly Pro	Leu Leu Arg Ser Ala	Leu Pro Ala Gly Trp	Phe Ile Ala			
1125	1130	1135				
gat aaa	tct gga gcc ggt gag	cgt ggg tct cgc ggt	atc att gca	6471		
Asp Lys	Ser Gly Ala Gly Glu	Arg Gly Ser Arg Gly	Ile Ile Ala			
1140	1145	1150				
gca ctg	ggg cca gat ggt aag	ccc tcc cgt atc gta	gtt atc tac	6516		
Ala Leu	Gly Pro Asp Gly Lys	Pro Ser Arg Ile Val	Val Ile Tyr			
1155	1160	1165				
acg acg	ggg agt cag gca act	atg gat gaa cga aat	aga cag atc	6561		
Thr Thr	Gly Ser Gln Ala Thr	Met Asp Glu Arg Asn	Arg Gln Ile			
1170	1175	1180				
gct gag	ata ggt gcc tca ctg	att aag cat tgg taactgtcag		6604		
Ala Glu	Ile Gly Ala Ser Leu	Ile Lys His Trp				
1185	1190					
accaagtttta	ctcatatata	ctttagattg	atttaaaaact	tcattttttaa	tttaaaagga	6664
tctaggtgaa	gataccttttt	gataatctca	tgacccaaaat	cccttaacgt	gagtttttcgt	6724
tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	ttcttgagat	ccttttttttc	6784
tgcgcgtaat	ctgctgcttg	caaacaaaaa	aaccaccgct	accagcgggtg	gtttgtttgc	6844
cggatcaaga	gctaccaact	ctttttccga	aggtaactgg	cttcagcaga	gcgcagatac	6904
caaatactgt	tcttctagt	tagccgtagt	taggccacca	cttcaagaac	tctgtagcac	6964
cgcctacata	cctcgctctg	ctaatacctgt	taccagtggc	tgctgccagt	ggcgataagt	7024
cgtgtcttac	cgggttggac	tcaagacgat	agttaccgga	taaggcgag	cggtcgggct	7084
gaacgggggg	ttcgtgcaca	cagcccagct	tgagcggaac	gacctacacc	gaactgagat	7144
acctacagcg	tgagctatga	gaaagcgcca	cgcttcccga	agggagaaaag	gcggacaggt	7204
atccggtaag	cggcagggtc	ggaacaggag	agcgcacgag	ggagcttcca	gggggaaacg	7264
cctgggtatct	ttatagtcct	gtcgggtttc	gccacctctg	acttgagcgt	cgattttttgt	7324
gatgctcgtc	agggggggcg	agcctatgga	aaaacgccag	caacgcggcc	tttttacggt	7384
tcctggcctt	ttgctggcct	tttgctcaca	tgttctttcc	tgcgttatcc	cctgattctg	7444
tggtataaccg	tattaccgcc	tttgagttag	ctgataccgc	tcgccgcagc	cgaacgaccg	7504

```

agcgcagcga gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc 7564
ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg 7624
gcagtgagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc caggctttac 7684
actttatgct tccggctcgt atgttggtgtg gaattgtgag cggataacaa tttcacacag 7744
gaaacagcta tgaccatgat tacgccagat ttaattaa 7782

```

```

<210> 44
<211> 120
<212> PRT
<213> Trình tự nhân tạo

```

```

<220>
<223> Cấu trúc tổng hợp

```

```

<400> 44

```

```

Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu
1           5           10           15

```

```

Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Ser Ser Ile Ser Asn Tyr
20           25           30

```

```

Tyr Trp Ser Trp Ile Arg Gln Ser Pro Gly Lys Gly Leu Glu Trp Ile
35           40           45

```

```

Gly Phe Ile Tyr Tyr Gly Gly Asn Thr Lys Tyr Asn Pro Ser Leu Lys
50           55           60

```

```

Ser Arg Val Thr Ile Ser Gln Asp Thr Ser Lys Ser Gln Val Ser Leu
65           70           75           80

```

```

Thr Met Ser Ser Val Thr Ala Ala Glu Ser Ala Val Tyr Phe Cys Ala
85           90           95

```

```

Arg Ala Ser Cys Ser Gly Gly Tyr Cys Ile Leu Asp Tyr Trp Gly Gln
100          105          110

```

```

Gly Thr Leu Val Thr Val Ser Ser
115          120

```

```

<210> 45

```

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

<220>
 <223> Cấu trúc tổng hợp

<400> 45

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
 1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
 20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
 35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
 50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
 65 70 75 80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
 85 90 95

Lys Val

<210> 46
 <211> 231
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 46

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
 1 5 10 15

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
 20 25 30

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
 35 40 45

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
 50 55 60

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
 65 70 75 80

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
 85 90 95

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
 100 105 110

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
 115 120 125

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
 130 135 140

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
 145 150 155 160

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
 165 170 175

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
 180 185 190

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
 195 200 205

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
 210 215 220

Ser Leu Ser Leu Ser Pro Gly
 225 230

<210> 47

<211> 129
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 47

Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val Val Gln Pro Gly Arg
 1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Thr Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ala Val Ile Ser Tyr Asp Ala Asn Tyr Lys Tyr Tyr Ala Asp Ser Val
 50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
 65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95

Ala Lys Asp Ser Gln Leu Arg Ser Leu Leu Tyr Phe Glu Trp Leu Ser
 100 105 110

Gln Gly Tyr Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
 115 120 125

Ser

<210> 48
 <211> 107
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Cấu trúc tổng hợp

<400> 48

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
65 70 75 80

Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
85 90 95

Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His
100 105

<210> 49

<211> 223

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Cấu trúc tổng hợp

<400> 49

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
1 5 10 15

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
20 25 30

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Glu Pro Asp
35 40 45

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
50 55 60

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

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
85 90 95

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
100 105 110

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
115 120 125

Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
130 135 140

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
145 150 155 160

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
165 170 175

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
180 185 190

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
195 200 205

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
210 215 220

<210> 50
<211> 286
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Cấu trúc tổng hợp

<400> 50

Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1 5 10 15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
 20 25 30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
 35 40 45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
 50 55 60

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

Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
 85 90 95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
 100 105 110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
 115 120 125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
 130 135 140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
 145 150 155 160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
 165 170 175

Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
 180 185 190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
 195 200 205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
 210 215 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser

225					230					235					240
Arg	Gly	Ile	Ile	Ala	Ala	Leu	Gly	Pro	Asp	Gly	Lys	Pro	Ser	Arg	Ile
				245					250					255	
Val	Val	Ile	Tyr	Thr	Thr	Gly	Ser	Gln	Ala	Thr	Met	Asp	Glu	Arg	Asn
			260					265					270		
Arg	Gln	Ile	Ala	Glu	Ile	Gly	Ala	Ser	Leu	Ile	Lys	His	Trp		
		275					280					285			