



(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)



1-0033766

(51)⁷ C07K 14/16

(13) B

(21) 1-2018-03005

(22) 15/12/2016

(86) PCT/EP2016/081159 15/12/2016

(87) WO 2017/102929 A1 22/06/2017

(30) 15200138.4 15/12/2015 EP; 16194124.0 17/10/2016 EP

(45) 25/10/2022 415

(43) 27/05/2019 374A

(73) Janssen Vaccines & Prevention B.V. (NL)

Archimedesweg 4, 2333 CN Leiden, Netherlands

(72) LANGEDIJK, Johannes, Petrus, Maria (NL); CALLENDRET, Benoit, Christophe, Stephan (NL); VAN MANEN, Danielle (NL); KRARUP, Anders (NL); STITZ, Jörn (NL); WEGMANN, Frank (NL); VELLINGA, Jort (NL).

(74) Công ty Cổ phần Sở hữu công nghiệp INVESTIP (INVESTIP)

(54) AXIT NUCLEIC MÃ HÓA PROTEIN VỎ HIV TỔNG HỢP, VECTƠ CHỨA AXIT NUCLEIC, CHẾ PHẨM VÀ VACXIN BAO GỒM CHẾ PHẨM NÀY

(57) Sáng chế đề cập đến axit nucleic mã hóa protein vỏ HIV tổng hợp, vectơ chứa axit nucleic, tế bào chứa vectơ, protein vỏ HIV tổng hợp, chế phẩm và vacxin bao gồm chế phẩm này. Vectơ biểu hiện virus mã hoá protein vỏ HIV tổng hợp có thể được sử dụng trong vacxin để mang lại sự miễn dịch bảo vệ được cải thiện chống lại sự nhiễm virus gây suy giảm miễn dịch ở người (human immunodeficiency virus - HIV).

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến axit nucleic mã hoá protein này, vector chứa axit nucleic, protein vỏ HIV tổng hợp và chế phẩm chứa protein này, và phương pháp kích thích miễn dịch bảo vệ kháng lại sự nhiễm virus gây suy giảm miễn dịch ở người (HIV). Vector biểu hiện virus mã hoá protein vỏ của HIV có thể được sử dụng trong vaccin để tạo ra đáp ứng miễn dịch bảo vệ được cải thiện kháng lại HIV.

Tình trạng kỹ thuật của sáng chế

Virus gây suy giảm miễn dịch ở người (Human Immunodeficiency Virus - HIV) ảnh hưởng đến hàng triệu người trên thế giới, và việc ngăn ngừa HIV thông qua vaccin hữu hiệu vẫn được ưu tiên rất cao, ngay cả trong thời đại điều trị kháng retrovirus phổ biến. HIV-1 là chủng virus gây bệnh phổ biến nhất, với hơn 90% ca HIV/AIDS có nguồn gốc từ sự nhiễm HIV-1 nhóm M. Nhóm M được chia tiếp thành các nhánh hoặc các kiểu phụ khác. Vaccin hữu hiệu lý tưởng là có khả năng gây ra cả đáp ứng tế bào hiệu nghiệm và kháng thể trung hòa rộng có khả năng làm trung hòa các chủng HIV-1 từ các nhánh khác nhau.

Mức độ đột biến cao của HIV-1 làm cho việc phát triển vaccin HIV-1 trở thành thách thức chưa từng có. Để cải thiện mức độ phủ của epitop tế bào T tiềm năng, và cải thiện đáp ứng tế bào, các kháng nguyên Gag, Pol và Env của HIV-1 “khảm”, có nguồn gốc từ các protein Kháng Nguyên Nhóm (Group Antigen - Gag), Polymeraza (Polymerase - Pol), và Vỏ (Envelope - Env) của HIV, được mô tả bởi các tác giả khác và được phát triển với nỗ lực mang lại mức độ phủ tối đa của epitop tế bào T tiềm năng (ví dụ như, Barouch et al, *Nat Med* 2010, 16: 319-323). Kháng nguyên khảm tương tự về chiều dài và cấu trúc miền với kháng nguyên HIV-1 có trong tự nhiên, kiểu đại.

Ví dụ, kháng nguyên HIV khảm được mô tả và được sử dụng trong vaccin bao gồm các kháng nguyên được mô tả trong tài liệu Barouch et al, *ở trên*; và WO 2010/059732 chẳng hạn như::

(a) kháng nguyên khảm Gag bao gồm:

(a)(i) trình tự Gag khảm thứ nhất (“mos1Gag”) có trình tự axit amin như được nêu trong SEQ ID NO: 1, và

(a)(ii) trình tự Gag khảm thứ hai (“mos2Gag”) có trình tự axit amin như được nêu trong SEQ ID NO: 2;

(b) kháng nguyên khảm Pol bao gồm:

(b)(i) trình tự Pol khảm thứ nhất (“mos1Pol”) có trình tự axit amin như được nêu trong SEQ ID NO: 3, và

(b)(ii) trình tự Pol khảm thứ hai (“mos2Pol”) có trình tự axit amin như được nêu trong SEQ ID NO: 4; và

(c) kháng nguyên khảm Env bao gồm:

(c)(i) trình tự Env khảm thứ nhất (“mos1Env”) có trình tự axit amin như được nêu trong SEQ ID NO: 5, và

(c)(ii) trình tự Env khảm thứ hai (“mos2Env”) có trình tự axit amin như được nêu trong SEQ ID NO: 6.

Trình tự mã hoá các kháng nguyên này được tách dòng trong vectơ, ví dụ, như vectơ adenovirut tái tổ hợp, ví dụ như, typ huyết thanh adenovirut tái tổ hợp 26 (rAd26), và các vectơ này được sử dụng trước đây làm vaccin để tạo ra đáp ứng miễn dịch đối với kháng nguyên (xem tài liệu ví dụ như Barouch et al, *ở trên*; và WO 2010/059732). Ví dụ, trình tự kháng nguyên khảm mos1Gag và mos1Pol thường được kết hợp vào protein dung hợp của Gag và Pol (“mos1GagPol”), và trình tự mã hóa của chúng được tách dòng vào vectơ Ad26 thứ nhất (“rAd26.mos1GagPol”); và các trình tự kháng nguyên mos2Gag và mos2Pol được kết hợp vào protein dung hợp khác của Gag và Pol (“mos2GagPol”), và trình tự mã hóa của chúng được tách dòng vào vectơ Ad26 thứ hai (“rAd26.mos2GagPol”). Cấu trúc mã hoá mos1Env và mos2Env thường được tách dòng vào vectơ Ad26 riêng rẽ (lần lượt là “rAd26.mos1Env” và “rAd26.mos2Env”).

Tập hợp của các kháng nguyên khảm này như được mô tả ở trên tạo ra mức độ phủ tổng thể tốt của thể phân lập HIV-1 Nhóm M, trong đó vectơ rAd26 mã hoá trình tự kháng nguyên 1 khảm (ví dụ như, rAd26.mos1GagPol và rAd26.mos1Env) thiên vị nhánh B và kiểu phụ HIV-1 CRF01, và vectơ rAd26 mã hoá trình tự kháng nguyên 2 khảm (ví dụ như, rAd26.mos2GagPol và rAd26.mos2Env) thiên vị chủng nhánh C. Các kháng nguyên Gag, Pol, và Env HIV-1 khảm được biểu hiện ở vectơ rAd26 có thể được sử dụng để cải thiện cả độ rộng và độ sâu của đáp ứng tế bào lympho T đặc hiệu kháng nguyên ở khỉ rhesus, mà không làm thỏa hiệp độ lớn của cả đáp ứng tế bào và thể dịch khi so với kháng nguyên HIV-1 trình tự liên ứng hoặc tự nhiên (Barouch et al, *ở trên*; và WO 2010/059732).

Tuy nhiên, khi tiếp tục nỗ lực phát triển thành phần vaccin được mô tả ở trên, đã phát hiện ra rằng rAd26.mos2Env thể hiện sự biểu hiện bề mặt tế bào và đáp ứng miễn dịch không tối ưu ở động vật linh trưởng không phải người, mà hơn nữa thể hiện độ ổn định di truyền không tối ưu, bất ngờ và không dự đoán trước được chưa được báo cáo cho đến nay trong quy trình sản xuất khi so với các vectơ rAd26 khác, như rAd26.mos1Env. Do đó, vaccin chứa rAd26.mos2Env có thể dẫn đến đáp ứng miễn

dịch không tối ưu chống lại kiểu phụ HIV-1 Nhánh C, vì kháng nguyên khả mos2Env thiên vị chủng HIV-1 nhánh C. Theo đó, có nhu cầu về phương án thay thế cho kháng nguyên mos2Env trong vacxin chống lại HIV mà có thể được sử dụng để gây ra đáp ứng miễn dịch được cải thiện chống lại HIV-1 nhánh C.

Bản chất kỹ thuật của sáng chế

Sáng chế đề cập đến protein vỏ virus gây suy giảm miễn dịch ở người (HIV) tổng hợp mới làm cải thiện sự biểu hiện bề mặt tế bào và độ ổn định di truyền khi so với kháng nguyên mos2Env được mô tả trước đó. Sáng chế cũng đề cập đến chế phẩm và phương pháp sử dụng protein vỏ HIV tổng hợp mới này và/hoặc trình tự mã hóa của chúng để gây ra đáp ứng miễn dịch chống lại HIV-1 tăng lên, cụ thể là HIV-1 nhánh C, tốt hơn là khi được sử dụng ở dạng kết hợp với các kháng nguyên HIV khác.

Theo một khía cạnh chung, sáng chế đề cập đến axit nucleic mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin của SEQ ID NO: 8, hoặc SEQ ID NO:8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (i) I529P (tức là, đột biến thay thế Ile cho Pro ở vị trí 529 của SEQ ID NO:8), (ii) K480E (tức là, đột biến thay thế Lys cho Glu ở vị trí 480 của SEQ ID NO:8), và (iii) sự kết hợp của EK479-480RRRR (tức là sự thay thế của Glu-Lys ở vị trí 479-480 của SEQ ID NO:8 bằng bốn gốc Arg liên tiếp), I529P, A471C (tức là, đột biến thay thế Ala cho Cys ở vị trí 471 của SEQ ID NO:8) và T575C (tức là, đột biến thay thế Thr cho Cys ở vị trí 575 của SEQ ID NO:8). Theo một phương án, protein vỏ HIV tổng hợp còn chứa trình tự tín hiệu, ví dụ trình tự tín hiệu có trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 9-12. Theo một phương án, trình tự tín hiệu có trình tự axit amin của SEQ ID NO: 9.

Theo các phương án nhất định, protein vỏ HIV tổng hợp còn chứa vùng xuyên màng, tốt hơn là vùng xuyên màng có trình tự axit amin được nêu trong SEQ ID NO: 13.

Theo các phương án nhất định, protein vỏ HIV tổng hợp còn chứa mảnh của vùng tế bào chất, tốt hơn là mảnh của vùng tế bào chất có chứa trình tự axit amin được nêu trong SEQ ID NO:14, hoặc các axit amin 1-4 của chúng (tức là, NRVR). Theo phương án trong đó protein vỏ HIV tổng hợp còn chứa vùng xuyên màng và mảnh của vùng tế bào chất, ưu tiên là protein cũng có chứa trình tự axit amin được nêu trong SEQ ID NO: 37, được dung hợp với đầu tận cùng carboxy (đầu tận cùng C) của SEQ ID NO:8 và đầu tận cùng amin (đầu tận cùng N) của vùng xuyên màng.

Theo phương án khác, protein vỏ HIV tổng hợp có chứa vùng trime hoá, ví dụ, vùng trime hoá có chứa trình tự axit amin được nêu trong SEQ ID NO: 15 (GCN4) hoặc SEQ ID NO:16 (miền foldon). Theo một phương án ưu tiên, vùng trime hoá có chứa trình tự axit amin được nêu trong SEQ ID NO: 15. Các phương án này với vùng trime

hoá hữu dụng để hòa tan (tức là không liên kết màng) protein vỏ HIV tổng hợp dựa trên trình tự vùng ngoại bào (ectodomain) được đề xuất trong bản mô tả này, chẳng hạn như có chứa trình tự axit amin được nêu trong SEQ ID NO: 8, trong đó vùng trime hoá nằm ở đầu tận cùng C của protein vỏ HIV tổng hợp.

Theo các phương án khác nữa, protein vỏ HIV tổng hợp có chứa SEQ ID NO: 8 với các đột biến sau: EK479-480RRRR, I529P, A471C và T575C. Việc đưa 6 gốc arginin liên tiếp (các vị trí 478 và 481 trong trình tự tự nhiên được nêu trong SEQ ID NO: 8 đã là gốc Arg) dẫn đến vị trí phân cắt furin được tối ưu hóa thêm, sao cho thu được vùng ngoại bào được xử lý (tức là được phân cắt) được cải thiện. Ba đột biến I529P, A471C và T575C được biết đến dưới dạng đột biến SOSIP, từ đó hai đột biến cuối cùng dẫn đến việc đưa vào cầu disulfua có thể có ở giữa các gốc xystein mới được tạo ra. Nhìn chung, các đột biến này dẫn đến protein vỏ HIV tổng hợp, được trime hóa, hòa tan mà không cần có vùng trime hoá.

Theo phương án được ưu tiên, sáng chế đề cập đến axit nucleic mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 17, SEQ ID NO: 18, hoặc aa 1-686 của SEQ ID NO: 19. Tốt nhất là protein vỏ HIV tổng hợp được mã hóa bởi axit nucleic có chứa hoặc gồm có trình tự axit amin được nêu trong SEQ ID NO: 18.

Theo một khía cạnh chung khác, sáng chế đề cập đến vector có chứa axit nucleic mã hoá protein vỏ HIV tổng hợp theo phương án của sáng chế. Theo một phương án, vector là vector virus. Theo phương án được ưu tiên, vector virus là vector adenovirus. Theo một phương án ưu tiên, vector adenovirus là vector adenovirus 26.

Khía cạnh chung khác của sáng chế đề cập đến chế phẩm, tốt hơn là chế phẩm vaccin, có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector theo phương án của sáng chế, và chất mang, trong đó axit nucleic mã hoá protein vỏ HIV tổng hợp được liên kết có điều khiển với trình tự khởi động. Theo một phương án, chế phẩm có chứa vector adenovirus, tốt hơn là vector adenovirus 26, mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18.

Theo một khía cạnh chung khác, sáng chế đề cập đến vaccin kết hợp để gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng cần chúng. Vaccin kết hợp này có chứa chế phẩm thứ nhất có chứa có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector, tốt hơn là vector adenovirus, tốt hơn nữa là vector adenovirus 26, mã hoá protein vỏ HIV tổng hợp có trình tự axit amin được nêu trong SEQ ID NO: 18, chế phẩm thứ hai có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector thứ hai, tốt hơn là vector adenovirus thứ hai, tốt hơn nữa là vector adenovirus 26 thứ hai, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin

được nêu trong SEQ ID NO: 5, và tùy ý ít nhất một chế phẩm bổ sung có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của ít nhất một vectơ được chọn từ nhóm gồm có vectơ mã hoá polypeptit kháng nguyên có trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 1-4, 28 và 29, và polypeptit có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của polypeptit kháng nguyên HIV được phân lập, bao gồm, nhưng không chỉ giới hạn ở, polypeptit có các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7, hoặc polypeptit có các gốc 30-724 của SEQ ID NO:36, trong đó chế phẩm thứ nhất, chế phẩm thứ hai và chế phẩm bổ sung tùy ý có mặt trong cùng chế phẩm hoặc trong một hoặc nhiều chế phẩm khác nhau.

Khía cạnh chung khác nữa của sáng chế đề cập đến phương pháp gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng cần chủng, có chứa bước dùng cho đối tượng chế phẩm hoặc vacxin kết hợp theo phương án của sáng chế. Sáng chế cũng đề cập đến phương pháp gây ra đáp ứng miễn dịch chống lại HIV có chứa bước môi và thúc đẩy đáp ứng miễn dịch bằng cách sử dụng chế phẩm hoặc vacxin kết hợp theo phương án của sáng chế.

Khía cạnh khác nữa của sáng chế đề cập đến protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8, hoặc SEQ ID NO: 8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (i) I529P, (ii) K480E, (iii) tổ hợp các đột biến EK479-480RRRR, I529P, A471C và T575C. Theo một phương án, protein vỏ HIV tổng hợp có chứa SEQ ID NO:8 có các đột biến EK479-480RRRR, I529P, A471C và T575C. Theo phương án khác, protein vỏ HIV tổng hợp có chứa các gốc 30-704 hoặc 30-711 của trình tự axit amin được nêu trong SEQ ID NO: 18. Theo phương án khác nữa protein vỏ HIV tổng hợp có chứa các gốc 30-686 của trình tự axit amin được nêu trong SEQ ID NO:19.

Khía cạnh khác của sáng chế đề cập đến tế bào, tốt hơn là tế bào được phân lập, có chứa vectơ theo phương án của sáng chế.

Mô tả vắn tắt các hình vẽ

Phần bản chất kỹ thuật của sáng chế nêu trên, cũng như là phần mô tả chi tiết sáng chế dưới đây, sẽ được hiểu rõ hơn khi tham khảo hình vẽ kèm theo. Cần hiểu rằng sáng chế không bị giới hạn ở các phương án chính xác được thể hiện trong hình vẽ.

FIG. 1A- 1C là hình vẽ dạng sơ đồ của cấu trúc của protein vỏ HIV; FIG. 1A thể hiện protein vỏ HIV có chiều dài đầy đủ; FIG. 1B thể hiện cấu trúc của protein vỏ HIV chuỗi đơn hòa tan theo phương án của sáng chế trong đó vùng xuyên màng (TM) được thay thế bằng vùng trime hoá GCN4, và vị trí phân cắt furin bị đột biến (sC4); FIG. 1C thể hiện cấu trúc của protein vỏ HIV liên kết màng theo phương án của sáng chế có chứa vùng xuyên màng và mảnh của vùng tế bào chất (C4D7);

FIG. 2 thể hiện mức độ biểu hiện của protein vỏ sC1 hòa tan, được dựa trên trình tự kháng nguyên khảm mos2Env với vùng trime hoá đầu tận cùng C bổ sung, và protein vỏ HIV tổng hợp hòa tan (sC4) theo phương án của sáng chế; sự biểu hiện được đo bằng thâm tách Western định lượng bằng cách sử dụng kháng thể đa dòng chống lại gp120; plasmit mã hoá sC1 hoặc sC4 được biểu hiện tạm thời hai lần, và mỗi chuyển nạp được định lượng hai lần bằng cách đo tỉ trọng; protein sC1 thể hiện mức độ biểu hiện rất thấp khi so với protein vỏ HIV tổng hợp sC4, mà thể hiện mức độ biểu hiện tương đối cao;

FIG. 3A và FIG. 3B thể hiện liên kết của protein vỏ HIV tổng hợp với kháng thể đơn dòng 17b (mAb17b) trong sự có mặt (màu xám nhạt) và sự vắng mặt (màu xám đậm) của CD4 hòa tan như xác định bằng thử nghiệm ELISA; FIG. 3A thể hiện liên kết của sC1; FIG. 3B thể hiện liên kết của sC4;

FIG. 4 là ảnh của thâm tách Western từ điện di gel polyacrylamit tự nhiên của protein sC1, và protein vỏ HIV tổng hợp sC4;

FIG. 5 thể hiện mức độ biểu hiện bề mặt tế bào tương đối của protein vỏ HIV tổng hợp C1, C1D7, C4 và C4D7 liên kết màng bằng phân tích FACS của tế bào biểu hiện các protein này bằng cách sử dụng kháng thể đa dòng kháng gp120 (GP120), và bằng cách liên kết với kháng thể trung hòa rộng PG9 (PG9) và PG16 (PG16) mà phụ thuộc cấu trúc bậc bốn và liên kết ưu tiên với trime Env đã gấp nếp đúng;

FIG. 6 là sự biểu diễn đồ họa của độ ổn định của vector adenovirut chứa trình tự mã hoá protein vỏ HIV tổng hợp theo sáng chế bao gồm C4 chiều dài đầy đủ (FLC4), C4D7, và sC4 sau nhiều lần cấy truyền virut; vector adenovirut tái tổ hợp 26 được tạo ra trong tế bào PER.C6; sau 3 lần cấy truyền ban đầu để chuyển nạp và tinh chế màng, 5 màng được chọn và được tăng quy mô cho 10 lần cấy truyền ở định dạng T25, dẫn đến tổng số lần cấy truyền virut (vpn) là 13; độ ổn định sau vpn 3, 5, 10, và 13 như xác định bằng phản ứng chuỗi polymeraza (PCR) bằng gen chuyển E1 được thể hiện; ví dụ, 3/5 có nghĩa là 3 màng ổn định trong số 5 màng được thử nghiệm, và 5/5 có nghĩa là 5 màng ổn định trong số 5 màng được thử nghiệm; và

FIG. 7A và FIG. 7B thể hiện độ chuẩn trung hòa virut chống lại hạt virut kiểu giả vỏ (EVP) HIV-1 trong thử nghiệm trung hòa dựa trên tế bào TZM-bl ở thỏ; giá trị IC_{50} được chuyển đổi log10 của nhóm dùng liều lượng vector adenovirut cao được đo chống lại EVP VSV-G (đối chứng âm) và MW965.26 (Bậc 1A nhánh C) ở các tuần 1, 8, 14, và 20; mỗi chấm biểu thị giá trị IC_{50} được chuyển đổi log10 của cá thể thỏ, với trung bình nhóm được chỉ ra bởi đường nằm ngang; HD: Độ Pha Loãng Cao Nhất được thử nghiệm (đường nét liền phía trên); LD: Độ Pha Loãng Thấp Nhất được thử nghiệm (đường nét liền phía dưới); LOB: giới hạn của nền, giá trị phần trăm 95 của mẫu âm được tuân theo (đường chấm chấm); Giá trị log10 IC_{50} vượt quá ngưỡng LD hoặc HD

được thiết lập ở đường tương ứng; so sánh không tham số một chiều với đối chứng bằng cách sử dụng phương pháp Dunn để xếp hạng chung được thực hiện cho mỗi thời điểm; sự khác biệt có ý nghĩa về mặt thống kê được chỉ ra trong biểu đồ: * = $P < 0,05$, ** = $P < 0,01$, và *** — $P < 0,001$; FIG. 7A thể hiện kết quả với VSV-G (đối chứng âm); và FIG. 7B thể hiện kết quả với MW965.26 (Bậc 1A nhánh C).

Mô tả chi tiết sáng chế

Các tài liệu công bố, bài báo và bằng sáng chế khác nhau được viện dẫn hoặc được mô tả trong tình trạng kỹ thuật và trong toàn bộ bản mô tả này; mỗi tài liệu tham khảo này được kết hợp ở đây nhằm mục đích tham khảo đến toàn bộ nội dung của nó. Thảo luận về tài liệu, hành động, nguyên liệu, dụng cụ, vật phẩm hoặc dạng tương tự được bao gồm trong bản mô tả này là nhằm mục đích cung cấp ngữ cảnh cho sáng chế. Thảo luận này không phải là sự thừa nhận rằng bất kỳ hoặc tất cả các vấn đề này tạo thành một phần của tình trạng kỹ thuật đối với sáng chế bất kỳ được bộc lộ hoặc được yêu cầu bảo hộ.

Trừ khi chỉ có dẫn khác, tất cả các thuật ngữ kỹ thuật và khoa học dùng trong bản mô tả này có cùng nghĩa như thường được hiểu bởi người có hiểu biết trung bình trong lĩnh vực mà sáng chế này thuộc về. Mặt khác, các thuật ngữ nhất định được dùng trong bản mô tả này có nghĩa như được nêu trong bản mô tả. Tất cả các bằng sáng chế, đơn sáng chế đã công bố và tài liệu công bố được viện dẫn trong bản mô tả này được kết hợp để tham khảo như thể được nêu đầy đủ trong bản mô tả này.

Như được dùng trong bản mô tả này, “đối tượng” có nghĩa là động vật bất kỳ, tốt hơn là động vật có vú, tốt nhất là người, mà sẽ được cho dùng hoặc đã được cho dùng vector, chế phẩm hoặc vaccin kết hợp theo phương án của sáng chế. Thuật ngữ “động vật có vú” như được dùng trong bản mô tả này, bao hàm động vật có vú bất kỳ. Ví dụ về động vật có vú bao gồm, nhưng không chỉ giới hạn ở, bò, ngựa, cừu, lợn, mèo, chó, chuột nhắt, chuột cống, thỏ, chuột lang, khỉ, con người, v.v., tốt hơn nữa là con người.

Sáng chế đề cập chung đến protein vỏ HIV tổng hợp, axit nucleic và vector mã hoá protein vỏ HIV tổng hợp này, và phương pháp gây ra đáp ứng miễn dịch chống lại HIV bằng vector mã hoá protein vỏ HIV tổng hợp hoặc protein vỏ HIV tổng hợp, một mình hoặc kết hợp với một hoặc nhiều vector bổ sung mã hoá một hoặc nhiều polypepit kháng nguyên HIV bổ sung và/hoặc kết hợp với một hoặc nhiều polypepit kháng nguyên HIV được phân lập bổ sung.

Virut gây suy giảm miễn dịch ở người (HIV) là thành viên của chi Lentivirinae, một phần của họ Retroviridae. Hai loài HIV lây nhiễm ở người: HIV-1 và HIV-2. HIV-1 là chủng phổ biến nhất của virut HIV, và được biết là gây bệnh nhiều hơn HIV-2. Như

được dùng trong bản mô tả này, thuật ngữ “virus gây suy giảm miễn dịch ở người” và “HIV” dùng để chỉ, nhưng không chỉ giới hạn ở, HIV-1 và HIV-2.

HIV được chia thành nhiều nhánh có mức độ sai khác di truyền cao. Như được dùng trong bản mô tả này, thuật ngữ “nhánh HIV” hoặc “kiểu phụ HIV” dùng để chỉ virus gây suy giảm miễn dịch ở người liên quan được phân loại theo mức độ tương tự di truyền của chúng. Hiện nay có ba nhóm thể phân lập HIV-1: M, N và O. Nhóm M (các chủng chính) gồm có ít nhất mười nhánh, từ A đến J. Nhóm O (các chủng phía ngoài) có thể gồm có số lượng nhánh tương tự. Nhóm N là thể phân lập HIV-1 mới mà chưa được phân loại vào nhóm M hoặc nhóm O.

Như được dùng trong bản mô tả này, thuật ngữ “polypeptit kháng nguyên HIV,” “protein kháng nguyên HIV,” và “chất sinh miễn dịch HIV” dùng để chỉ polypeptit có khả năng gây ra đáp ứng miễn dịch, ví dụ như, đáp ứng thể dịch và/hoặc qua trung gian tế bào, chống lại HIV ở đối tượng. Polypeptit kháng nguyên có thể là protein của HIV, mảnh hoặc epitop của chúng, hoặc dạng kết hợp của nhiều protein HIV hoặc phần của chúng, có thể gây ra đáp ứng miễn dịch hoặc tạo ra sự miễn dịch, ví dụ như, sự miễn dịch bảo vệ, chống lại HIV ở đối tượng.

Tốt hơn là, polypeptit kháng nguyên có khả năng làm tăng đáp ứng miễn dịch bảo vệ ở vật chủ, ví dụ như, gây ra đáp ứng miễn dịch chống lại bệnh virus hoặc sự nhiễm virus, và/hoặc tạo ra sự miễn dịch ở (tức là, chủng ngừa) đối tượng chống lại bệnh virus hoặc sự nhiễm virus, mà bảo vệ đối tượng chống lại bệnh virus hoặc sự nhiễm virus. Ví dụ, polypeptit kháng nguyên có thể có chứa protein hoặc mảnh của chúng từ Virus Gây Thiếu Hụt Miễn Dịch Ở Khỉ (Simian Immunodeficiency Virus - SIV) hoặc HIV, như protein gp160 vỏ HIV hoặc SIV, protein chất nền/capsit HIV hoặc SIV, và sản phẩm gen *gag*, *pol* và *env* HIV hoặc SIV.

Polypeptit kháng nguyên HIV có thể là kháng nguyên HIV-1 hoặc HIV-2 bất kỳ hoặc mảnh của chúng. Ví dụ về kháng nguyên HIV bao gồm, nhưng không chỉ giới hạn ở sản phẩm gen *gag*, *pol*, và *env*, mà mã hoá protein cấu trúc và enzym thiết yếu. Sản phẩm gen *gag*, *pol*, và *env* được tổng hợp như polyprotein, được xử lý thêm thành nhiều sản phẩm protein khác. Sản phẩm protein sơ cấp của gen *gag* là polyprotein gag protein cấu trúc virus, được xử lý thêm thành các sản phẩm protein MA, CA, SP1, NC, SP2, và P6. Gen *pol* mã hoá enzym virus (Pol, polymeraza), và sản phẩm protein sơ cấp được xử lý thêm thành các sản phẩm protein RT, ARNaza H, IN, và PR. Gen *env* mã hoá protein cấu trúc, cụ thể là glycoprotein của vỏ hạt virus. Sản phẩm protein sơ cấp của gen *env* là gp160, được xử lý thêm thành gp120 và gp41. Ví dụ khác về kháng nguyên HIV bao gồm các protein điều hòa gen Tat và Rev; các protein phụ trợ Nef, Vpr, Vif và Vpu; protein capsit, protein nucleocapsit, và protein virus p24.

Theo các phương án nhất định, polypeptit kháng nguyên HIV có chứa kháng nguyên Gag, Env, hoặc Pol HIV, hoặc phần kháng nguyên hoặc epitop bất kỳ hoặc dạng kết hợp của chúng, tốt hơn là kháng nguyên Gag, Env, hoặc Pol HIV-1 hoặc phần kháng nguyên bất kỳ hoặc epitop hoặc dạng kết hợp của chúng.

Polypeptit kháng nguyên HIV cũng có thể là kháng nguyên HIV khảm. Như được dùng trong bản mô tả này, “kháng nguyên khảm” dùng để chỉ protein tái tổ hợp được lắp ráp từ mảnh có trình tự tự nhiên. Kháng nguyên khảm giống với kháng nguyên tự nhiên, nhưng được tối ưu hóa để tối đa hóa mức độ phủ của epitop tế bào T tiềm năng được tìm thấy trong trình tự tự nhiên, mà cải thiện độ rộng và độ bao phủ của đáp ứng miễn dịch. Kháng nguyên HIV khảm được sử dụng theo sáng chế tốt hơn là kháng nguyên Gag, Pol, và/hoặc Env khảm, và tốt hơn nữa là kháng nguyên Gag, Pol, và/hoặc Env HIV-1 khảm. Như được dùng trong bản mô tả này, “kháng nguyên Gag, Pol, và/hoặc Env HIV khảm” cụ thể là dùng để chỉ kháng nguyên khảm có chứa nhiều epitop có nguồn gốc từ một hoặc nhiều trình tự polyprotein Gag, Pol và/hoặc Env của HIV.

Theo một phương án, kháng nguyên HIV khảm được sử dụng theo sáng chế là kháng nguyên Gag HIV khảm với epitop có nguồn gốc từ trình tự của sản phẩm gen *gag* (các ví dụ được nêu trong các SEQ ID NO: 1, 2); kháng nguyên Pol HIV khảm với epitop có nguồn gốc từ trình tự của sản phẩm gen *pol* (các ví dụ được nêu trong các SEQ ID NO: 3, 4); hoặc kháng nguyên Env HIV khảm với epitop có nguồn gốc từ trình tự của sản phẩm gen *env* (các ví dụ được nêu trong các SEQ ID NO: 5, 6; cũng như là kháng nguyên mới của sáng chế, ví dụ như trong các SEQ ID NO: 8, 17, 18, 19, có thể được coi là kháng nguyên Env HIV khảm). Theo các phương án nhất định, kháng nguyên HIV khảm được sử dụng theo sáng chế có thể có chứa dạng kết hợp của các epitop có nguồn gốc từ các trình tự của các sản phẩm gen *gag*, *pol*, và/hoặc *env*. Các ví dụ minh họa và không làm giới hạn sáng chế bao gồm các kháng nguyên Env-Pol khảm với các epitop có nguồn gốc từ các trình tự của sản phẩm gen *env* và *pol*; các kháng nguyên Gag-Pol khảm với các epitop có nguồn gốc từ các trình tự của sản phẩm gen *gag* và *pol*; và các kháng nguyên Gag-Env khảm với các epitop có nguồn gốc từ các trình tự của sản phẩm gen *gag* và *env*. Trình tự của sản phẩm gen *gag*, *pol*, và *env* có thể là có nguồn gốc từ một hoặc nhiều nhánh.

Ví dụ về kháng nguyên Gag, Pol và/hoặc Env HIV khảm có thể được sử dụng theo sáng chế bao gồm các kháng nguyên được mô tả trong, ví dụ như, US20120076812; Barouch et al., *Nat Med* 2010, 16:319-323; và Barouch et al., *Cell* 155:1-9, 2013, tất cả các tài liệu này được kết hợp trong bản mô tả này chỉ để tham khảo. Tốt hơn là, kháng nguyên Gag, Pol, và/hoặc Env HIV khảm được sử dụng theo sáng chế bao gồm, nhưng không chỉ giới hạn ở, mos1Gag (SEQ ID NO: 1), mos2Gag (SEQ ID NO: 2), mos1Pol

(SEQ ID NO: 3), mos2Pol (SEQ ID NO: 4), mos1Env (SEQ ID NO: 5), mos2Env (SEQ ID NO: 6), mos1GagPol (SEQ ID NO: 28), mos2GagPol (SEQ ID NO: 29), và dạng kết hợp của chúng.

Như được dùng trong bản mô tả này, thuật ngữ “protein vỏ HIV,” “protein Env,” và “Env” dùng để chỉ protein được biểu hiện trên vỏ của hạt virus HIV và làm cho HIV nhắm đích và gắn vào màng huyết tương của tế bào nhiễm HIV, hoặc mảnh hoặc dẫn xuất của chúng mà có thể gây ra đáp ứng miễn dịch hoặc tạo ra sự miễn dịch chống lại HIV ở đối tượng cần chúng. Gen *env* HIV mã hoá protein tiền chất gp160, được phân cắt bằng cách phân giải protein thành hai glycoprotein vỏ trưởng thành, gp120 và gp41. Phản ứng phân cắt được điều khiển bằng proteaza tế bào chủ, furin, ở trình tự được bảo toàn cao trong tiền chất glycoprotein vỏ retrovirus. Cụ thể hơn, gp160 trimer hóa thành (gp160)₃ và sau đó trải qua phân cắt thành hai phần được kết hợp không cộng hóa trị gp120 và gp41. Sự xâm nhập virus sau đó được điều khiển bằng trimer của đimer khác loại gp120/gp41. Gp120 là mảnh liên kết thụ thể, và liên kết với thụ thể CD4 trên tế bào đích mà có thụ thể này, như, ví dụ như, tế bào T-trợ giúp. Gp41, được liên kết không cộng hóa trị với gp120, là mảnh dung hợp và cung cấp bước thứ hai nhờ đó HIV đi vào tế bào. Gp41 ban đầu được che giấu trong vỏ virus, nhưng khi gp120 liên kết với thụ thể CD4, gp120 thay đổi hình dạng của nó làm cho gp41 để trở nên được lộ ra, khi đó nó có thể giúp đỡ trong việc dung hợp với tế bào chủ. Gp140 là vùng ngoại bào không được phân cắt của gp160 trimer, tức là, (gp160)₃, đã được sử dụng làm dạng thay thế cho trạng thái tự nhiên của gai virus, được phân cắt.

Theo phương án của sáng chế, “protein vỏ HIV” có thể là protein gp160, gp140, gp120, gp41, dạng kết hợp, dung hợp, dạng cắt cụt hoặc dẫn xuất của chúng. Ví dụ, “protein vỏ HIV” có thể bao gồm protein gp120 được kết hợp không cộng hóa trị với protein gp41. Nó cũng có thể bao gồm protein gp140 trimer được làm ổn định mà có thể có hoặc có thể được cải biến để bao gồm vùng trimer hoá mà làm ổn định trimer của gp140. Ví dụ về vùng trimer hoá bao gồm, nhưng không chỉ giới hạn ở, vùng trimer hoá “foldon” T4-fibrin; vùng trimer hoá cuộn xoắn có nguồn gốc từ GCN4; và đơn vị con xúc tác của *E. coli* aspartat transcarbamoylaza làm đuôi trimer. “Protein vỏ HIV” cũng có thể là protein vỏ HIV bị cắt cụt bao gồm, nhưng không giới hạn, protein vỏ có chứa phần cắt cụt đầu tận cùng C trong vùng ngoại bào (tức là miền mà kéo dài vào không gian ngoại bào), phần cắt cụt trong gp41, chẳng hạn như phần cắt cụt trong vùng xuyên màng của gp41, hoặc phần cắt cụt trong vùng tế bào chất của gp41. “Protein vỏ HIV” có thể còn là dẫn xuất của protein vỏ HIV có trong tự nhiên có đột biến trình tự, ví dụ như, ở vị trí phân cắt furin, và/hoặc còn được gọi là đột biến SOSIP.

Tốt hơn là, “protein vỏ HIV” là “protein vỏ HIV tổng hợp.” Như được dùng trong bản mô tả này, thuật ngữ “protein vỏ HIV tổng hợp” dùng để chỉ protein vỏ HIV không có trong tự nhiên được tối ưu hóa để gây ra đáp ứng miễn dịch hoặc tạo ra sự miễn dịch chống lại một hoặc nhiều chủng HIV có trong tự nhiên ở đối tượng cần chúng. Protein Env HIV khả năng là ví dụ về protein Env HIV tổng hợp, và sáng chế đề xuất kháng nguyên Env HIV tổng hợp mới, ví dụ như kháng nguyên Env HIV tổng hợp có chứa SEQ ID NO: 8, 17, 18, hoặc 19.

Protein vỏ HIV tổng hợp và trình tự mã hóa của chúng

Phương án của sáng chế đề cập đến protein vỏ HIV tổng hợp mới và phân tử axit nucleic mã hoá các protein này.

Theo một phương án, sáng chế đề cập đến protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8, hoặc SEQ ID NO:8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (i) I529P, (ii) K480E, và (iii) tổ hợp các đột biến EK479- 480RRRR, I529P, A471C và T575C. SEQ ID NO:8 có chứa gp120 trưởng thành tổng hợp và gp41 bị cắt cụt tổng hợp mà không có vùng xuyên màng, hay vùng tế bào chất. SEQ ID NO:8 là trình tự không có trong tự nhiên có chứa thể khảm của trình tự từ kháng nguyên khảm mos2Env (SEQ ID NO: 6), và trình tự protein vỏ HIV khác. Trình tự của kháng nguyên Env tổng hợp mới có chứa SEQ ID NO:8 được tối ưu hóa để tạo ra độ phủ rộng và đáp ứng tế bào T được tăng cường chống lại HIV nhánh C (khi so với kháng nguyên mos2Env (SEQ ID NO: 6)). Theo các phương án nhất định, các axit amin khác có thể được bổ sung vào SEQ ID NO: 8 hoặc một trong các biến thể của nó được xác định trong bản mô tả này.

Theo các phương án nhất định, protein vỏ HIV tổng hợp còn chứa trình tự tín hiệu. Protein vỏ HIV tổng hợp được tổng hợp với trình tự tín hiệu được phân cắt từ chuỗi polypeptit mới sinh trong quá trình vận chuyển nó vào lumen của lưới nội chất (endoplasmic reticulum - ER). Về nguyên tắc, trình tự tín hiệu đã biết bất kỳ có thể được sử dụng. Tốt hơn là trình tự tín hiệu Env HIV hoặc biến thể của chúng được sử dụng. Các trình tự tín hiệu khác nhau đã được sử dụng trong lĩnh vực kỹ thuật này cho protein Env HIV (xem tài liệu ví dụ như WO 2014/107744). Theo các phương án nhất định, trình tự tín hiệu có chứa SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11 hoặc SEQ ID NO:12. Theo một phương án ưu tiên, trình tự tín hiệu có chứa SEQ ID NO: 9.

Theo các phương án nhất định, protein vỏ HIV tổng hợp còn chứa vùng xuyên màng. Vùng xuyên màng neo protein vỏ HIV tổng hợp vào màng ER, và góp phần vào việc lắp ráp màng và chức năng của vỏ HIV. Tốt hơn là, vùng xuyên màng có chứa SEQ ID NO:13.

Theo phương án khác, protein vỏ HIV tổng hợp có chứa gp41 có vùng tế bào chất bị cắt cụt. gp41 có vùng tế bào chất dài bất thường ở đầu cacboxyl của nó, thường là khoảng 150 axit amin (Edwards et al., *J. Virology*, 2002, 76:2683-2691). Phần cắt cụt của vùng tế bào chất được báo cáo là gây ra sự bộc lộ của vùng được bảo toàn trong vùng ngoại bào của protein Env HIV-1 (*Id.*). Vùng tế bào chất bị cắt cụt trong vỏ HIV tổng hợp theo sáng chế nằm trong khoảng từ một đến khoảng 140 axit amin, chẳng hạn như 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, hoặc 140 axit amin của vùng tế bào chất chiều dài đầy đủ. Theo các phương án nhất định vùng tế bào chất bị cắt cụt có nguồn gốc từ các axit amin 704-862 của SEQ ID NO: 17 (tức là từ vùng tế bào chất của phân tử C4 theo sáng chế), bằng sự cắt cụt sau axit amin nhất định đến đầu tận cùng C. Theo phương án được ưu tiên, protein vỏ HIV tổng hợp có chứa vùng tế bào chất bị cắt cụt có từ 1 đến 10 gốc axit amin, tốt hơn là từ 4 đến 8 gốc axit amin, và tốt nhất là 7 gốc axit amin của vùng tế bào chất gp41 HIV. Vùng tế bào chất hoặc mảnh của chúng của protein vỏ HIV tổng hợp nằm ở từ đầu tận cùng C đến vùng ngoại bào (vùng ngoại bào-ectodomain), và khi protein vỏ HIV tổng hợp cũng có chứa vùng xuyên màng, vùng tế bào chất hoặc mảnh của chúng nằm ở từ đầu tận cùng C đến vùng xuyên màng. Xem, ví dụ như, FIG. 1A và FIG. 1C. Theo phương án cụ thể, protein vỏ HIV tổng hợp có chứa gp41 với vùng tế bào chất bị cắt cụt có trình tự axit amin được nêu trong SEQ ID NO:14 hoặc mảnh của chúng, chẳng hạn như các gốc 1-4 của chúng (tức là NRVR). Vùng tế bào chất bị cắt cụt khác đã được mô tả và có thể được sử dụng (ví dụ như Schiernle et al., *PNAS* 1997; Abrahamyan et al., *J Virol* 2005).

Theo phương án trong đó protein vỏ HIV tổng hợp còn chứa vùng xuyên màng và mảnh của vùng tế bào chất, ưu tiên là protein này cũng có chứa trình tự axit amin được nêu trong SEQ ID NO: 37, mà chứa các gốc 655-682 của SEQ ID NO: 18, trong đó trình tự axit amin được nêu trong SEQ ID NO: 37 được dung hợp với đầu tận cùng C của SEQ ID NO: 8 và đầu tận cùng N của vùng xuyên màng.

Theo phương án được đặc biệt ưu tiên của sáng chế, protein vỏ HIV tổng hợp còn chứa vùng xuyên màng, chẳng hạn như vùng xuyên màng có trình tự axit amin được nêu trong SEQ ID NO:13, và vùng tế bào chất bị cắt cụt hoặc mảnh của vùng tế bào chất, chẳng hạn như mảnh tế bào chất có trình tự axit amin được nêu trong SEQ ID NO: 14 hoặc các gốc 1-4 của SEQ ID NO:14 (tức là, NRVR). Tốt nhất là, protein vỏ HIV tổng hợp có chứa hoặc gồm có trình tự axit amin được nêu trong SEQ ID NO: 18, có hoặc không có trình tự tín hiệu (tức là, các axit amin 1-29 của SEQ ID NO:18).

Theo phương án khác, protein vỏ HIV tổng hợp có chứa vùng trime hoá mà thay thế vùng xuyên màng Env. Vùng trime hoá làm tăng độ ổn định của cấu trúc trime Env.

Tốt hơn là, protein vỏ HIV tổng hợp có chứa polypeptit gp140 được cải biến để bao gồm vùng trime hoá mà làm ổn định trime của gp140. Ví dụ về vùng trime hoá bao gồm, nhưng không chỉ giới hạn ở, vùng trime hoá “foldon” T4-fibritin, chẳng hạn như vùng trime hoá có chứa trình tự axit amin của SEQ ID:16; vùng trime hoá cuộn xoắn có nguồn gốc từ GCN4, chẳng hạn như vùng trime hoá có chứa trình tự axit amin của SEQ ID:15; đơn vị con xúc tác của *E. coli* aspartat transcarbamoylaza làm đuôi trime; hoặc môtip trime hóa dựa trên matrillin. Nếu có mặt, vùng trime hoá thường nằm ở từ đầu tận cùng C đến vùng ngoại bào (xem FIG. 1B). Theo phương án được ưu tiên nhất định trong đó protein vỏ HIV tổng hợp có chứa vùng trime hoá, protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 19, có hoặc không có trình tự tín hiệu (tức là, các gốc axit amin 1-29 của SEQ ID NO:19). Các phương án này với vùng trime hoá chủ yếu hữu dụng đối với biến thể vùng ngoại bào hòa tan của protein vỏ HIV tổng hợp. Theo các phương án nhất định của các biến thể hòa tan này của sáng chế, có thể gây đột biến vị trí phân cắt furin (ví dụ như đột biến của Lys cho Glu ở vị trí 480 trong SEQ ID NO: 8) để làm bất hoạt vị trí phân cắt này, sao cho protein sẽ là chuỗi đơn; nó kết hợp tốt với vùng trime hoá, đặc biệt là với vùng trime hoá GCN4 của SEQ ID NO: 19.

Phiên bản thay thế của các biến thể vùng ngoại bào hòa tan này của protein vỏ HIV tổng hợp mà không sử dụng vùng trime hoá cũng là phương án của sáng chế, và có thể được điều chế từ SEQ ID NO: 8 bằng cách kết hợp các đột biến mà làm tối ưu hóa vị trí phân cắt furin (thay thế dipeptit Gly-Lys ở vị trí 479-480 bằng 4 gốc Arg) cũng như là đột biến được gọi là SOSIP (đột biến I>P ở vị trí 529, và đưa cầu disulfua vào giữa các vị trí 471 và 575 bằng cách thay thế mỗi Ala và Thr tương ứng trên các vị trí này trong SEQ ID NO: 8 bằng gốc Cys). Điều này tạo ra protein có trình tự axit amin được nêu trong SEQ ID NO: 8 với dạng kết hợp sau đây của các đột biến: EK479-480RRRR, I529P, A471C và T575C.

Một cải biến có thể có để làm tăng hơn nữa hàm lượng trime của protein vỏ HIV tổng hợp theo sáng chế (có chứa SEQ ID NO: 8), là sự cải biến của Ile thành Pro ở vị trí 529. Điều này có thể hữu hiệu đối với cả biến thể hòa tan và biến thể liên kết màng.

Vector

Khía cạnh chung khác của sáng chế đề cập đến vector có chứa axit nucleic mã hoá protein vỏ HIV tổng hợp. Theo phương án của sáng chế, vector có thể có chứa protein vỏ HIV tổng hợp bất kỳ được mô tả trong bản mô tả này. Theo phương án được ưu tiên của sáng chế, vector có chứa axit nucleic mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8, SEQ ID NO:17, SEQ ID NO: 18, hoặc SEQ ID NO: 19, và tốt hơn là SEQ ID NO: 18.

Theo phương án của sáng chế, axit nucleic mã hoá protein vỏ HIV tổng hợp được liên kết có điều khiển với vùng khởi động, có nghĩa là axit nucleic này chịu sự kiểm soát của vùng khởi động. Vùng khởi động có thể là vùng khởi động cùng loại (tức là, có nguồn gốc từ cùng nguồn di truyền với vector) hoặc vùng khởi động khác loại (tức là, có nguồn gốc từ vector khác hoặc nguồn di truyền khác). Ví dụ về vùng khởi động thích hợp bao gồm vùng khởi động virus cự bào (CMV) và vùng khởi động virus gây ung thư mô liên kết Rous (RSV). Tốt hơn là, vùng khởi động nằm ngược chiều với axit nucleic ở trong băng biểu hiện. Trình tự khởi động CMV ví dụ mà có thể được liên kết có điều khiển với axit nucleic mã hoá protein vỏ HIV tổng hợp được thể hiện trong SEQ ID NO: 24.

Theo phương án của sáng chế, vector có thể là vector biểu hiện. Vector biểu hiện bao gồm, nhưng không chỉ giới hạn ở, vector để biểu hiện protein tái tổ hợp và vector để chuyển axit nucleic vào đối tượng để biểu hiện ở mô của đối tượng, chẳng hạn như vector virus. Ví dụ về vector virus thích hợp được sử dụng theo sáng chế bao gồm, nhưng không chỉ giới hạn ở vector adenovirus, vector virus kết hợp adeno, vector virus đậu mùa, vector MVA, vector virus ruột, vector virus gây viêm não ngựa Venezuela, vector virus rừng Semliki, vector virus khảm thuốc lá, vector lentivirus, v.v.. Vector cũng có thể là vector không phải là virus. Ví dụ về vector không phải là virus bao gồm, nhưng không chỉ giới hạn ở plasmid, nhiễm sắc thể nhân tạo của vi khuẩn, nhiễm sắc thể nhân tạo của nấm men, thể thực khuẩn, v.v..

Theo các phương án nhất định của sáng chế, vector là vector adenovirus. Adenovirus theo sáng chế thuộc vào họ Adenoviridae, và tốt hơn là thuộc chi Mastadenovirus. Nó có thể là adenovirus người, nhưng cũng là adenovirus nhiễm các loài khác, bao gồm nhưng không giới hạn adenovirus bò (ví dụ như adenovirus bò 3, BAdV3), adenovirus chó (ví dụ như CAdV2), adenovirus lợn (ví dụ như PAdV3 hoặc 5), hoặc adenovirus khỉ (mà bao gồm adenovirus khỉ và adenovirus khỉ không đuôi, chẳng hạn như adenovirus tinh tinh hoặc adenovirus khỉ đột). Tốt hơn là, adenovirus là adenovirus người (HAdV, hoặc AdHu), hoặc adenovirus khỉ như adenovirus tinh tinh hoặc adenovirus khỉ đột (ChAd, AdCh, hoặc SAdV). Theo sáng chế, nó có nghĩa là adenovirus người nếu được đề cập đến dưới dạng Ad mà không chỉ thị loài, ví dụ như ký hiệu ngắn “Ad26” có cùng nghĩa với HAdV26, mà là typ huyết thanh adenovirus người 26. Cũng như được dùng trong bản mô tả này, ký hiệu “rAd” có nghĩa là adenovirus tái tổ hợp, ví dụ như, “rAd26” dùng để chỉ adenovirus người tái tổ hợp 26.

Các nghiên cứu tiên tiến nhất đã được thực hiện bằng cách sử dụng adenovirus người, và adenovirus người được ưu tiên theo các khía cạnh nhất định của sáng chế. Theo phương án được ưu tiên nhất định, adenovirus tái tổ hợp theo sáng chế được dựa

trên adenovirut người. Theo phương án được ưu tiên, adenovirut tái tổ hợp được dựa trên các typ huyết thanh adenovirut người 5, 11, 26, 34, 35, 48, 49, 50, 52, v.v. Theo phương án được đặc biệt ưu tiên của sáng chế, adenovirut là adenovirut người có typ huyết thanh 26. Ưu điểm của typ huyết thanh này là tỉ lệ có mặt trong huyết thanh thấp và/hoặc độ chuẩn kháng thể trung hòa tồn tại từ trước thấp trong quần thể người, và kinh nghiệm sử dụng ở đối tượng người trong thử nghiệm lâm sàng.

Adenovirut khi thường cũng có tỉ lệ có mặt trong huyết thanh thấp và/hoặc độ chuẩn kháng thể trung hòa tồn tại từ trước thấp trong quần thể người, và lượng đáng kể của công việc đã được báo cáo bằng cách sử dụng vectơ adenovirut tinh tinh (ví dụ như US6083716; WO 2005/071093; WO 2010/086189; WO 2010085984; Farina *et al*, 2001, *J Virol* 75: 11603-13 [13]; Cohen *et al*, 2002, *J Gen Virol* 83: 151-55 [69]; Kobinger *et al*, 2006, *Virology* 346: 394-401 [70]; Tatsis *et al.*, 2007, *Molecular Therapy* 15: 608-17 [71]; cũng xem phần xem xét của Bangari and Mittal, 2006, *Vaccine* 24: 849-62 [72]; và phần xem xét của Lasaro and Ertl, 2009, *Mol Ther* 17: 1333-39 [73]). Do đó, theo các phương án khác, adenovirut tái tổ hợp theo sáng chế được dựa trên adenovirut khi, ví dụ như adenovirut tinh tinh. Theo các phương án nhất định, adenovirut tái tổ hợp được dựa trên typ adenovirut khi 1, 7, 8, 21, 22, 23, 24, 25, 26, 27.1, 28.1, 29, 30, 31.1, 32, 33, 34, 35.1, 36, 37.2, 39, 40.1, 41.1, 42.1, 43, 44, 45, 46, 48, 49, 50 hoặc SA7P.

Tốt hơn là, vectơ adenovirut là vectơ virut tái tổ hợp thiếu khả năng sao chép, như rAd26, rAd35, rAd48, rAd5HVR48, v.v.

Theo phương án được ưu tiên của sáng chế, vectơ adenovirut có chứa protein capsit từ các typ huyết thanh hiếm bao gồm Ad26. Theo phương án điển hình, vectơ là virut rAd26. “Protein capsit adenovirut” dùng để chỉ protein trên capsit của adenovirut (ví dụ như, vectơ Ad26, Ad35, rAd48, rAd5HVR48) tham gia vào việc xác định typ huyết thanh và/hoặc tính hướng kích thích của adenovirut cụ thể. Protein capsit adenovirut thường bao gồm các protein sợi, penton và/hoặc hexon. Như được dùng trong bản mô tả này “protein capsit” đối với adenovirut cụ thể, chẳng hạn như “protein capsit Ad26” có thể là, ví dụ, protein capsit khảm mà bao gồm ít nhất một phần của protein capsit Ad26. Theo các phương án nhất định, protein capsit là protein capsit nguyên vẹn của Ad26. Theo các phương án nhất định, hexon, penton và sợi là của Ad26.

Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này sẽ nhận ra rằng các thành phần có nguồn gốc từ nhiều typ huyết thanh có thể được kết hợp trong vectơ adenovirut tái tổ hợp đơn lẻ. Do đó, adenovirut khảm mà kết hợp các tính chất mong muốn từ các typ huyết thanh khác có thể được tạo ra. Do đó, theo một số phương án, adenovirut khảm theo sáng chế có thể kết hợp sự không có mặt của sự miễn dịch tồn tại

từ trước của typ huyết thanh thứ nhất với các đặc điểm chẳng hạn như độ ổn định nhiệt độ, sự lắp ráp, sự neo, hiệu suất sản xuất, sự nhiễm được định hướng lại hoặc được cải thiện, độ ổn định của ADN trong tế bào đích, và dạng tương tự.

Theo các phương án nhất định vectơ adenovirut tái tổ hợp hữu dụng theo sáng chế có nguồn gốc chủ yếu hoặc toàn bộ từ Ad26 (*tức là*, vectơ là rAd26). Theo một số phương án, adenovirut thiếu khả năng sao chép, ví dụ như, bởi vì nó chứa sự làm khuyết trong vùng E1 của hệ gen. Đối với adenovirut có nguồn gốc từ adenovirut không phải nhóm C, như Ad26 hoặc Ad35, nó thường trao đổi trình tự mã hóa E4-orf6 của adenovirut với E4- orf6 của adenovirut của nhóm phụ C ở người như Ad5. Điều này cho phép sự nhân giống của các adenovirut này trong các dòng tế bào bổ sung đã biết rõ mà biểu hiện gen E1 của Ad5, chẳng hạn như ví dụ tế bào 293, tế bào PER.C6, và dạng tương tự (*xem, ví dụ* Havenga, et al., 2006, *J Gen Virol* 87: 2135-43; WO 03/104467). Tuy nhiên, các adenovirut này sẽ không có khả năng sao chép ở tế bào không bổ thể mà không biểu hiện gen E1 của Ad5.

Điều chế vectơ adenovirut tái tổ hợp đã được biết rõ trong lĩnh vực kỹ thuật này. Điều chế vectơ rAd26 được mô tả, ví dụ như, trong WO 2007/104792 và trong Abbink *et al.*, (2007) *Virol* 81(9): 4654-63. Trình tự hệ gen ví dụ của Ad26 được tìm thấy trong Mã số truy cập GenBank EF 153474 và trong SEQ ID NO:1 của WO 2007/104792. Ví dụ về vectơ hữu dụng cho sáng chế ví dụ bao gồm các vectơ được mô tả trong WO2012/082918, phần bộc lộ của chúng được kết hợp ở đây chỉ để tham khảo.

Thường là, vectơ hữu dụng theo sáng chế được tạo ra bằng cách sử dụng axit nucleic có chứa toàn bộ hệ gen adenovirut tái tổ hợp (*ví dụ như*, vectơ plasmit, cosmit, hoặc baculovirut). Do đó, sáng chế cũng đề xuất phân tử axit nucleic được phân lập mã hoá vectơ adenovirut theo sáng chế. Phân tử axit nucleic theo sáng chế có thể ở dạng ARN hoặc ở dạng ADN thu được bằng cách tách dòng hoặc sản xuất bằng cách tổng hợp. ADN có thể là sợi kép hoặc sợi đơn.

Vectơ adenovirut hữu dụng theo sáng chế thường thiếu khả năng sao chép. Theo các phương án này, virut được làm cho thiếu khả năng sao chép bằng cách làm khuyết hoặc làm bất hoạt các vùng thiết yếu để sao chép của virut, chẳng hạn như vùng E1. Vùng có thể là về cơ bản bị làm khuyết hoặc làm bất hoạt bằng cách, ví dụ, cài xen gen quan tâm, chẳng hạn như gen mã hoá protein vỏ HIV tổng hợp (thường liên kết với vùng khởi động), hoặc gen mã hoá polypeptit kháng nguyên HIV (thường liên kết với vùng khởi động) ở trong vùng này. Theo một số phương án, vectơ theo sáng chế có thể chứa sự làm khuyết ở các vùng khác, chẳng hạn như các vùng E2, E3 hoặc E4, hoặc sự cài xen của gen khác loại liên kết với vùng khởi động ở trong một hoặc nhiều vùng này. Đối với adenovirut được đột biến E2 và/hoặc E4, thông thường dòng tế bào bổ sung E2

và/hoặc E4 được sử dụng để tạo ra adenovirut tái tổ hợp. Đột biến ở vùng E3 của adenovirut không cần được bổ sung bởi dòng tế bào, vì E3 không cần thiết để sao chép.

Dòng tế bào đóng gói thường được sử dụng để tạo ra lượng đủ của vector adenovirut để sử dụng theo sáng chế. Tế bào đóng gói là tế bào mà có chứa các gen mà đã bị làm khuyết hoặc làm bất hoạt ở vector thiếu khả năng sao chép, do đó cho phép virut sao chép trong tế bào. Tế bào đóng gói thích hợp đối với adenovirut có sự làm khuyết trong vùng E1 bao gồm, ví dụ, PER.C6, 911, 293, và E1 A549.

Theo phương án của sáng chế, và như lưu ý ở trên, protein vỏ HIV tổng hợp bất kỳ được mô tả trong bản mô tả này có thể được biểu hiện ở vector theo sáng chế. Căn cứ vào sự thoái hóa của mã di truyền, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này hiểu rõ rằng một vài trình tự axit nucleic có thể được thiết kế mà mã hoá cùng protein, theo phương pháp hoàn toàn thông thường trong lĩnh vực kỹ thuật này. Axit nucleic mã hoá protein vỏ HIV tổng hợp có thể tùy ý được tối ưu hóa bộ ba mã hóa để đảm bảo sự biểu hiện phù hợp ở vật chủ được xử lý (*ví dụ như*, người). Sự tối ưu hóa bộ ba mã hóa là công nghệ được áp dụng rộng rãi trong lĩnh vực kỹ thuật này. Một số ví dụ không làm giới hạn sáng chế về trình tự mã hoá protein vỏ HIV tổng hợp theo sáng chế được nêu trong các SEQ ID NO: 25, 26 và 27. Thông thường, axit nucleic mã hoá protein vỏ HIV tổng hợp được tách dòng vào vùng E1 và/hoặc E3 của hệ gen adenovirut.

Theo phương án được ưu tiên của sáng chế, vector là vector adenovirut, và tốt hơn là vector rAd26, tốt nhất là vector rAd26 với ít nhất một sự làm khuyết trong vùng E1 của hệ gen adenovirut, ví dụ như chẳng hạn như được mô tả trong Abbink, *J Virol*, 2007. 81(9): p. 4654-63, được kết hợp trong bản mô tả này để tham khảo.

Sáng chế còn đề xuất tế bào, tốt hơn là tế bào được phân lập, có chứa vector bất kỳ được mô tả trong bản mô tả này. Tế bào có thể được sử dụng cho sự sản xuất protein tái tổ hợp, hoặc cho sự sản xuất hạt virut.

Do đó các phương án của sáng chế cũng liên quan đến phương pháp tạo ra polypeptit kháng nguyên HIV tổng hợp. Phương pháp này có chứa bước chuyển nạp tế bào chủ bằng vector biểu hiện có chứa axit nucleic mã hoá polypeptit kháng nguyên HIV tổng hợp liên kết có điều khiển với vùng khởi động, cho sinh trưởng tế bào đã được chuyển nạp trong điều kiện thích hợp để biểu hiện polypeptit kháng nguyên HIV tổng hợp, và phân lập polypeptit kháng nguyên HIV tổng hợp từ tế bào. Polypeptit kháng nguyên HIV tổng hợp có thể được phân lập hoặc được thu gom từ tế bào bằng phương pháp bất kỳ đã biết trong lĩnh vực kỹ thuật này bao gồm sắc ký ái lực, v.v. Kỹ thuật được sử dụng để biểu hiện protein tái tổ hợp được biết rõ đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này dựa vào sáng chế.

Sáng chế cũng bao gồm phương pháp sản xuất vectơ mã hoá polypeptit kháng nguyên HIV tổng hợp theo sáng chế, phương pháp này có chứa bước nuôi cấy tế bào mà có chứa vectơ, để nhân giống và nhân lên vectơ trong quá trình nuôi cấy, và phân lập vectơ mà mã hoá polypeptit kháng nguyên HIV tổng hợp theo sáng chế từ môi trường nuôi cấy tế bào, ví dụ như từ tế bào, từ môi trường nuôi cấy, hoặc từ cả hai. Vectơ có thể được tinh chế thêm theo các phương pháp đã biết trong lĩnh vực kỹ thuật này.

Theo các phương án nhất định, sáng chế đề xuất vectơ theo phương án của sáng chế có chứa axit nucleic mã hoá polypeptit kháng nguyên HIV tổng hợp, và theo phương án ví dụ nhất định axit nucleic này có trình tự nucleotit được chọn từ nhóm gồm có SEQ ID NO: 25, 26 và 27.

Chế phẩm

Theo một khía cạnh chung khác, sáng chế đề cập đến chế phẩm có chứa vectơ có chứa axit nucleic mã hoá protein vỏ HIV tổng hợp và chất mang. Theo phương án của sáng chế, vectơ bất kỳ được mô tả trong bản mô tả này có thể được bao gồm trong chế phẩm. Tốt hơn là, vectơ là vectơ virut, tốt hơn là vectơ adenovirut, và tốt hơn nữa là vectơ adenovirut 26. Theo phương án được ưu tiên, chế phẩm có chứa vectơ adenovirut, tốt hơn là vectơ adenovirut 26 mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8, SEQ ID NO: 18, hoặc SEQ ID NO: 19, và tốt hơn nữa là trình tự axit amin được nêu trong SEQ ID NO: 18.

Theo một khía cạnh, sáng chế đề xuất vaccin kết hợp có chứa một hoặc nhiều vectơ cùng nhau có chứa trình tự axit nucleic mã hoá (i) protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 (ví dụ như SEQ ID NO: 18 hoặc 19) và (ii) protein vỏ HIV thứ hai có chứa trình tự axit amin được nêu trong SEQ ID NO: 5. Mỗi vectơ có thể ở trong chế phẩm riêng rẽ, hoặc được kết hợp trong chế phẩm đơn lẻ. Cả hai axit nucleic trong (các) vectơ được dự định là được sử dụng cho một đối tượng, mà sẽ dẫn đến đáp ứng miễn dịch với HIV mà rộng hơn so với đáp ứng miễn dịch mà sẽ thu được khi sử dụng một trong hai vectơ một mình. Cả hai trình tự axit nucleic có thể cùng có mặt trên một vectơ đơn lẻ.

Theo phương án của sáng chế, chế phẩm có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vectơ, chẳng hạn như vectơ virut. Như được dùng trong bản mô tả này, “lượng hữu hiệu để gây đáp ứng miễn dịch” hoặc “lượng hữu hiệu về mặt miễn dịch” có nghĩa là lượng chế phẩm đủ để gây ra tác dụng miễn dịch hoặc đáp ứng miễn dịch mong muốn ở đối tượng cần chúng. Theo một phương án, lượng hữu hiệu để gây đáp ứng miễn dịch có nghĩa là lượng đủ để gây ra đáp ứng miễn dịch ở đối tượng cần chúng. Theo phương án khác, lượng hữu hiệu để gây đáp ứng miễn dịch có nghĩa là lượng đủ

để tạo ra sự miễn dịch ở đối tượng cần chúng, ví dụ như, mang lại tác dụng bảo vệ chống lại bệnh chẳng hạn như sự nhiễm virus. Lượng hữu hiệu để gây đáp ứng miễn dịch có thể thay đổi tùy thuộc vào nhiều yếu tố, chẳng hạn như tình trạng sức khỏe của đối tượng, tuổi, khối lượng, sức khỏe, v.v.; ứng dụng cụ thể, có gây ra đáp ứng miễn dịch hay tạo ra sự miễn dịch bảo vệ hay không; vectơ tái tổ hợp cụ thể được dùng; polypeptit gây miễn dịch hoặc polypeptit kháng nguyên được mã hóa bởi vectơ tái tổ hợp được dùng; polypeptit kháng nguyên cụ thể được dùng; và bệnh cụ thể, ví dụ như, sự nhiễm virus, mà mong muốn có sự miễn dịch này đối với nó. Lượng hữu hiệu để gây đáp ứng miễn dịch có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này dựa vào sáng chế.

Chỉ dẫn chung là, lượng hữu hiệu để gây đáp ứng miễn dịch khi được sử dụng có tham chiếu đến vectơ virus tái tổ hợp chẳng hạn như vectơ adenovirus có thể nằm trong khoảng từ khoảng 10^8 hạt virus đến khoảng 10^{12} hạt virus, ví dụ như 10^8 , 10^9 , 10^{10} , 10^{11} , hoặc 10^{12} hạt virus. Lượng hữu hiệu để gây đáp ứng miễn dịch có thể được dùng trong chế phẩm đơn lẻ, hoặc trong nhiều chế phẩm, như 1, 2, 3, 4, 5, 6, 7, 8, 9, hoặc 10 chế phẩm (ví dụ như, viên nén, viên nang hoặc thuốc tiêm), trong đó việc sử dụng nhiều viên nang hoặc việc tiêm cùng nhau cung cấp cho đối tượng với lượng hữu hiệu để gây đáp ứng miễn dịch. Nhìn chung, khi được sử dụng có tham chiếu đến polypeptit, chẳng hạn như polypeptit kháng nguyên được phân lập, lượng hữu hiệu để gây đáp ứng miễn dịch nằm trong khoảng từ, ví dụ như từ khoảng 0,3 đến khoảng 3000 microgam (μg), ví dụ như 1-1000 μg , ví dụ như 10-500 μg , ví dụ như khoảng 10, 50, 100, 150, 200, 250, 300, 350, 400, 450 hoặc 500 μg . Ví dụ không làm giới hạn sáng chế là, có thể kết hợp việc sử dụng vectơ mã hoá kháng nguyên Env HIV tổng hợp theo sáng chế (có SEQ ID NO: 8) với việc sử dụng polypeptit Env, ví dụ như 250 μg protein trime Env nhánh C HIV có các axit amin 30-708 của SEQ ID NO: 7. Cũng có thể sử dụng lượng hữu hiệu để gây đáp ứng miễn dịch cho đối tượng, và sau đó sử dụng liều lượng khác của lượng hữu hiệu để gây đáp ứng miễn dịch cho cùng đối tượng, trong chế độ trị liệu gọi là môi-thúc đẩy. Ý tưởng chung của chế độ trị liệu môi-thúc đẩy được biết rõ đối với người có hiểu biết trung bình trong lĩnh vực vaccin. Các lần sử dụng thúc đẩy khác có thể tùy ý được bổ sung vào chế độ trị liệu, khi cần.

Chế phẩm theo sáng chế còn chứa chất mang. Chất mang có thể bao gồm một hoặc nhiều tá dược được dùng như chất gắn, chất phân rã, tác nhân làm phòng, chất tạo huyền phù, chất nhũ tương hóa, chất làm ẩm, chất bôi trơn, hương liệu, chất làm ngọt, chất bảo quản, thuốc nhuộm, chất làm tan và chất phủ. Bản chất chính xác của chất mang hoặc nguyên liệu khác có thể phụ thuộc vào đường dùng, ví dụ như, các đường trong cơ, dưới da, qua đường miệng, trong tĩnh mạch, qua da, trong niêm mạc (ví dụ

như, ruột), trong mũi hoặc trong màng bụng. Đối với chế phẩm tiêm được dạng lỏng, ví dụ, huyền phù và dung dịch, chất mang và chất phụ trợ thích hợp bao gồm nước, glycol, dầu, rượu, chất bảo quản, chất tạo màu và chất tương tự. Đối với chế phẩm dùng qua đường miệng dạng rắn, ví dụ, bột, viên nang, viên nén hình elip, gelcap và viên nén, chất mang và chất phụ trợ thích hợp bao gồm tinh bột, đường, chất pha loãng, chất tạo hạt, chất bôi trơn, chất gắn, chất làm phân rã và chất tương tự. Đối với hỗn hợp chất phun/thuốc xông qua đường mũi, dung dịch/huyền phù trong nước có thể có chứa nước, glycol, dầu, chất làm mềm, chất ổn định, chất làm ẩm, chất bảo quản, chất thơm, hương vị, và chất tương tự làm chất mang và chất phụ trợ thích hợp.

Chế phẩm theo sáng chế có thể được tạo chế phẩm trong vật chất bất kỳ thích hợp để sử dụng cho đối tượng để làm thuận lợi cho việc sử dụng và cải thiện hiệu quả, bao gồm, nhưng không giới hạn ở, việc sử dụng qua đường miệng (trong ruột) và tiêm ngoài đường tiêu hóa. Việc tiêm ngoài đường tiêu hóa bao gồm tiêm hoặc truyền trong tĩnh mạch, tiêm trong động mạch, tiêm dưới da, tiêm trong cơ, và tiêm trong khớp. Chế phẩm theo sáng chế cũng có thể được tạo chế phẩm cho các đường dùng khác bao gồm qua niêm mạc, qua mắt, qua trực tràng, cấy ghép tác động lâu dài, sử dụng dưới lưỡi, dưới lưỡi, từ niêm mạc miệng bỏ qua sự tuần hoàn tĩnh mạch, xông, hoặc trong mũi.

Theo phương án nhất định của sáng chế, chế phẩm có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vectơ adenovirut được tinh chế hoặc được tinh chế một phần, chẳng hạn như vectơ adenovirut 26, có chứa axit nucleic mã hoá protein vỏ HIV tổng hợp theo sáng chế. Chế phẩm này có thể được tạo chế phẩm dưới dạng vaccin (cũng được đề cập đến dưới dạng “chế phẩm gây miễn dịch”) theo phương pháp đã biết rõ trong lĩnh vực.

Chế phẩm theo sáng chế có thể còn tùy ý có chứa tá dược để tăng cường đáp ứng miễn dịch. Thuật ngữ "chất bổ trợ" và "chất kích thích miễn dịch" được sử dụng thay thế lẫn nhau trong bản mô tả này, và được định nghĩa là một hoặc nhiều chất gây kích thích hệ miễn dịch. Trong ngữ cảnh này, tá dược được sử dụng để tăng cường đáp ứng miễn dịch đối với vectơ mã hoá protein vỏ HIV tổng hợp theo sáng chế và/hoặc polypeptit kháng nguyên HIV được sử dụng ở dạng kết hợp với vectơ mã hoá protein vỏ HIV tổng hợp theo sáng chế.

Tá dược thích hợp được sử dụng theo sáng chế nên là tá dược mà có khả năng là an toàn, dung chịu tốt và hữu hiệu ở người, chẳng hạn như ví dụ QS-21, Detox-PC, MPL-SE, MoGM-CSF, TiterMax-G, CRL- 1005, GERBU, TERamit, PSC97B, Adjuver, PG-026, GSK-I, GcMAF, B-aethine, MPC-026, Adjuvax, CpG ODN, Betafectin, muối nhôm (ví dụ như AdjuPhos), Adjuplex, và MF59. Tỷ lệ tối ưu của mỗi

thành phần trong chế phẩm có thể được xác định bằng các kỹ thuật đã biết rõ đối với người có hiểu biết trung bình trong lĩnh vực dựa vào sáng chế.

Theo phương án được ưu tiên, tá dược là muối nhôm, chẳng hạn như AdjuPhos.

Việc điều chế và sử dụng chế phẩm gây miễn dịch đã được biết rõ đối với người có hiểu biết trung bình trong lĩnh vực. Dược phẩm lỏng thường bao gồm chất mang lỏng chẳng hạn như nước, dầu mỡ, dầu động vật hoặc thực vật, dầu khoáng hoặc dầu tổng hợp. Dung dịch nước muối sinh lý, dextroza hoặc dung dịch sacarit khác hoặc glycol chẳng hạn như etylen glycol, propylen glycol hoặc polyetylen glycol cũng có thể được bao gồm.

Ví dụ, vector adenovirut tái tổ hợp có thể được bảo quản trong chất đệm mà cũng được sử dụng đối với Adenovirut theo Tiêu chuẩn thế giới (Hoganson *et al.*, 2002, *Bioprocessing J* 1: 43-8): 20 mM Tris độ pH 8, 25 mM NaCl, 2,5% glyxerol. Chất đệm chế phẩm hữu dụng khác thích hợp để dùng cho người là 20 mM Tris, 2 mM MgCl₂, 25 mM NaCl, sucroza 10% khối lượng/thể tích, polysorbat-80 0,02% khối lượng/thể tích. Chất đệm chế phẩm khác thích hợp cho adenovirut tái tổ hợp có chứa 10-25 mM chất đệm xitrat độ pH 5,9-6,2, 4-6% (khối lượng/khối lượng) hydroxypropyl-beta-cyclodextrin (HBCD), 70-100 mM NaCl, 0,018-0,035% (khối lượng/khối lượng) polysorbat-80, và tùy ý 0,3-0,45% (khối lượng/khối lượng) etanol. Tất nhiên, nhiều chất đệm khác có thể được sử dụng, và một số ví dụ về chế phẩm thích hợp để bảo quản và để sử dụng các vector đã được trình bày trong dược phẩm là đã được biết.

Theo phương án của sáng chế, chế phẩm theo sáng chế có thể được sử dụng cùng với một hoặc nhiều vector bổ sung mã hoá một hoặc nhiều polypepit kháng nguyên HIV bổ sung, và/hoặc một hoặc nhiều polypeptit kháng nguyên HIV được phân lập. Vector bổ sung và/hoặc polypeptit kháng nguyên HIV có thể có mặt trong cùng chế phẩm có chứa protein Env HIV tổng hợp theo sáng chế. Chúng cũng có thể có mặt trong một hoặc nhiều chế phẩm khác nhau mà có thể được sử dụng cùng với chế phẩm có chứa protein Env HIV tổng hợp theo sáng chế trong vaccin kết hợp. Tốt hơn là, một hoặc nhiều vector bổ sung là vector virut, chẳng hạn như vector adenovirut, và tốt nhất là vector adenovirut 26. Một hoặc nhiều vector bổ sung có thể mã hoá polypeptit kháng nguyên HIV bất kỳ đã được biết đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này dựa vào sáng chế.

Theo một phương án, chế phẩm hoặc vaccin kết hợp còn chứa vector adenovirut thứ hai, tốt hơn là vector adenovirut 26, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 5. Ưu điểm của các phương án này là độ rộng được tăng lên của đáp ứng miễn dịch (bao phủ các chủng từ Nhánh B và Nhánh C).

Theo phương án khác, chế phẩm hoặc vaccin kết hợp theo sáng chế còn chứa vectơ adenovirut, tốt hơn là vectơ adenovirut 26, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 28 (mos1GagPol).

Theo phương án khác, chế phẩm hoặc vaccin kết hợp theo sáng chế còn chứa vectơ adenovirut, tốt hơn là vectơ adenovirut 26, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 29 (mos2GagPol).

Theo phương án cụ thể, chế phẩm hoặc vaccin kết hợp theo sáng chế còn chứa vectơ adenovirut thứ hai, tốt hơn là vectơ adenovirut 26, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 5, và một hoặc nhiều vectơ adenovirut bổ sung, tốt hơn là vectơ adenovirut 26, mã hoá một hoặc nhiều polypeptit kháng nguyên HIV có chứa trình tự axit amin được chọn từ nhóm gồm có SEQ ID NO: 28 hoặc SEQ ID NO: 29. Ví dụ, chế phẩm hoặc vaccin kết hợp theo phương án của sáng chế có thể có chứa bốn vectơ adenovirut, tốt hơn là vectơ adenovirut 26, với vectơ thứ nhất mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 (ví dụ như SEQ ID NO: 18); vectơ thứ hai mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 5; vectơ thứ ba mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 28; và vectơ thứ tư mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 29.

Theo một số phương án, chế phẩm hoặc vaccin kết hợp còn chứa một hoặc nhiều polypeptit kháng nguyên HIV được phân lập. Polypeptit kháng nguyên HIV bất kỳ đã biết rõ đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này dựa vào sáng chế có thể cũng được bao gồm trong chế phẩm hoặc vaccin kết hợp theo sáng chế, bao gồm, nhưng không giới hạn protein vỏ HIV (ví dụ như, gp160, gp140, gp120, hoặc gp41), tốt hơn là protein gp140 trime được làm ổn định, như protein gp140 nhánh C hoặc nhánh A được làm ổn định. Theo phương án được ưu tiên, polypeptit kháng nguyên HIV được phân lập là protein gp140 trime nhánh C HIV được làm ổn định, chẳng hạn như protein gp140 trime nhánh C HIV được làm ổn định có chứa các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO:7 (các gốc 1-29 của SEQ ID NO:7 trong trình tự tín hiệu). Polypeptit Env HIV thay thế hoặc bổ sung mà có thể được sử dụng cùng với protein gp140 nhánh C hoặc một mình, là protein trime Env khảm, ví dụ có trình tự axit amin như được bộc lộ trong các axit amin 30-724 của SEQ ID NO: 36 (tương ứng với SEQ ID NO: 2 của WO 2014/107744, các gốc 1-29 của SEQ ID NO:36 trong trình tự tín hiệu).

Theo phương án cụ thể của sáng chế, protein kháng nguyên HIV có thể là protein vỏ HIV tổng hợp theo sáng chế. Do đó, protein vỏ tổng hợp theo sáng chế có thể được

sử dụng ở dạng được phân lập và/hoặc được tinh chế để gây ra đáp ứng miễn dịch hoặc mang lại sự miễn dịch bảo vệ, v.v. chống lại HIV ở đối tượng cần chúng. Protein vỏ tổng hợp bất kỳ được mô tả trong bản mô tả này có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 có thể được sử dụng làm protein kháng nguyên HIV ở dạng được phân lập và/hoặc được tinh chế. Theo phương án được ưu tiên, khi được sử dụng ở dạng được phân lập làm protein kháng nguyên HIV, protein vỏ tổng hợp có chứa các gốc 30-711 của trình tự axit amin được nêu trong SEQ ID NO: 18 hoặc các gốc 30-686 của trình tự axit amin được nêu trong SEQ ID NO: 19, và tốt hơn nữa là gốc các 30-704 của trình tự axit amin được nêu trong SEQ ID NO: 18. Polypeptit kháng nguyên HIV được phân lập cũng có thể có chứa SEQ ID NO: 8 với các đột biến sau: EK479-480RRRR, I529P, A471C và T575C.

Phương án của sáng chế cũng đề cập đến chế phẩm hoặc vaccin kết hợp có chứa protein vỏ HIV tổng hợp được phân lập có chứa trình tự axit amin được nêu trong SEQ ID NO: 8. Protein vỏ HIV tổng hợp bất kỳ được mô tả trong bản mô tả này có thể được sử dụng. Theo phương án cụ thể của sáng chế, protein vỏ HIV tổng hợp được phân lập có chứa các gốc 30-704 hoặc 30-711 của trình tự axit amin được nêu trong SEQ ID NO: 18, các gốc 30-686 của trình tự axit amin được nêu trong SEQ ID NO: 19, hoặc trình tự axit amin được nêu trong SEQ ID NO: 8 với các đột biến sau: EK479-480RRRR, I529P, A471C và T575C. Các chế phẩm hoặc vaccin kết hợp này có thể còn chứa một hoặc nhiều vectơ biểu hiện, ví dụ như, vectơ adenovirut chẳng hạn như vectơ adenovirut 26, mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung, chẳng hạn như protein vỏ HIV tổng hợp theo sáng chế, hoặc protein kháng nguyên HIV khác như được nêu trong các SEQ ID NO: 4, 5, 7, 28 hoặc 29, hoặc mảnh của chúng.

Sáng chế cũng đề cập đến phương pháp sản xuất chế phẩm hoặc vaccin kết hợp theo sáng chế. Theo phương án của sáng chế, phương pháp sản xuất chế phẩm hoặc chế phẩm kết hợp có chứa bước kết hợp vectơ có chứa axit nucleic mã hoá protein vỏ HIV tổng hợp theo sáng chế với chất mang, và tùy ý một hoặc nhiều vectơ bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung và/hoặc một hoặc nhiều polypeptit kháng nguyên HIV được phân lập. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này quen thuộc với các kỹ thuật thông thường được sử dụng để điều chế các chế phẩm này.

Vaccin Và Vaccin Kết Hợp

Theo phương án của sáng chế, chế phẩm có thể là vaccin. Như được dùng trong bản mô tả này, thuật ngữ “vaccin” dùng để chỉ chế phẩm có chứa vectơ biểu hiện, tốt hơn là vectơ virut, mã hoá protein vỏ HIV tổng hợp theo sáng chế mà có thể mang lại sự miễn dịch bảo vệ hoặc đáp ứng miễn dịch bảo vệ cho đối tượng, hoặc để chủng ngừa

đối tượng. Theo phương án của sáng chế, khi sử dụng chế phẩm cho đối tượng, vectơ biểu hiện biểu hiện protein vỏ HIV tổng hợp được mã hóa, và protein vỏ HIV tổng hợp được biểu hiện này trình diện với hệ miễn dịch của đối tượng, bằng cách đó gây ra đáp ứng yêu cầu để tạo ra sự miễn dịch, hoặc gây ra đáp ứng miễn dịch.

Do đó, theo một khía cạnh chung khác, sáng chế đề xuất vacxin để gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng. Theo phương án của sáng chế, vacxin có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vectơ biểu hiện mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8, và tốt hơn là trình tự axit amin được nêu trong SEQ ID NO:18. Tốt hơn là, vectơ biểu hiện là vectơ virus, tốt hơn nữa là vectơ adenovirus, và tốt nhất là vectơ adenovirus 26.

Theo phương án của sáng chế, “gây ra đáp ứng miễn dịch” khi được sử dụng để đề cập đến phương pháp và chế phẩm được mô tả trong bản mô tả này bao hàm việc mang lại sự miễn dịch bảo vệ và/hoặc chủng ngừa đối tượng chống lại sự nhiễm, như sự nhiễm HIV, nhằm mục đích phòng ngừa, cũng như là gây ra đáp ứng miễn dịch hoặc tác dụng mong muốn ở đối tượng cần chủng chống lại sự nhiễm, như sự nhiễm HIV, nhằm mục đích trị liệu. Tốt hơn là, phương pháp theo sáng chế là để nhằm mục đích phòng ngừa, như để mang lại sự miễn dịch bảo vệ. Đáp ứng miễn dịch có thể là đáp ứng miễn dịch tế bào và/hoặc đáp ứng miễn dịch thể dịch.

Như được dùng trong bản mô tả này, thuật ngữ “sự miễn dịch bảo vệ” hoặc “đáp ứng miễn dịch bảo vệ” có nghĩa là đối tượng chủng ngừa có thể kiểm soát sự nhiễm bởi tác nhân gây bệnh mà việc chủng ngừa được thực hiện để chống lại chúng. Thông thường, đối tượng đã phát triển “đáp ứng miễn dịch bảo vệ” chỉ phát triển triệu chứng lâm sàng từ nhẹ đến trung bình hoặc không có triệu chứng nào. Thông thường, đối tượng có “đáp ứng miễn dịch bảo vệ” hoặc “sự miễn dịch bảo vệ” chống lại tác nhân nhất định sẽ không tử vong do sự nhiễm tác nhân đó.

Theo phương án của sáng chế, chế phẩm vacxin có thể còn chứa một hoặc nhiều vectơ bổ sung, ví dụ như, vectơ virus, như vectơ adenovirus, tốt hơn là vectơ adenovirus 26, mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung và/hoặc một hoặc nhiều polypeptit kháng nguyên HIV được phân lập. Protein vỏ HIV tổng hợp, vectơ bổ sung và/hoặc một hoặc nhiều polypeptit kháng nguyên HIV được phân lập có thể được tạo chế phẩm trong cùng chế phẩm hoặc một hoặc nhiều chế phẩm khác nhau trong vacxin.

Sáng chế cũng đề cập đến vacxin kết hợp để mời và thúc đẩy đáp ứng miễn dịch đối với một hoặc nhiều nhánh HIV ở đối tượng cần chủng bằng cách sử dụng một hoặc nhiều vectơ ở dạng kết hợp với polypeptit kháng nguyên được phân lập. Do đó, theo

một khía cạnh chung khác, sáng chế đề xuất vaccin kết hợp để gây ra đáp ứng miễn dịch chống lại HIV ở đối tượng. Theo phương án của sáng chế, vaccin kết hợp có chứa:

(i) chế phẩm thứ nhất có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector biểu hiện mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 hoặc SEQ ID NO: 8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (a) I529P, (b) K480E, và (c) tổ hợp các đột biến EK479-480RRRR, I529P, A471C và T575C, và chất mang; và

(ii) chế phẩm thứ hai có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của polypeptit kháng nguyên HIV được phân lập và chất mang,

trong đó một trong số chế phẩm thứ nhất và chế phẩm thứ hai là để tạo môi đáp ứng miễn dịch và chế phẩm còn lại là để thúc đẩy đáp ứng miễn dịch.

Theo phương án của sáng chế, vaccin kết hợp tùy ý còn chứa lượng hữu hiệu để gây đáp ứng miễn dịch của một hoặc nhiều vector biểu hiện bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung. Một hoặc nhiều vector biểu hiện bổ sung có thể được bao gồm trong chế phẩm thứ nhất hoặc chế phẩm thứ hai, hoặc một hoặc nhiều vector biểu hiện bổ sung có thể được bao gồm trong một hoặc nhiều chế phẩm bổ sung để được dùng cùng với chế phẩm thứ nhất và/hoặc chế phẩm thứ hai.

Như được dùng trong bản mô tả này, thuật ngữ “cùng phân phối”, “cùng sử dụng” hoặc “được dùng cùng với” dùng để chỉ việc sử dụng đồng thời hai hoặc nhiều thành phần, chẳng hạn như vector biểu hiện virus và polypeptit kháng nguyên được phân lập, hoặc nhiều vector biểu hiện virus. “Việc sử dụng đồng thời” có thể là việc sử dụng hai hoặc nhiều thành phần ít nhất là trong cùng một ngày. Khi hai thành phần được “dùng cùng nhau,” chúng có thể được dùng trong các chế phẩm riêng rẽ một cách tuần tự trong khoảng thời gian ngắn, chẳng hạn như 24, 20, 16, 12, 8 hoặc 4 giờ, hoặc trong 1 giờ hoặc ngắn hơn, hoặc chúng có thể được dùng trong chế phẩm đơn lẻ vào cùng thời điểm.

Theo phương án cụ thể của vaccin kết hợp theo sáng chế, chế phẩm thứ nhất có chứa vector adenovirus, tốt hơn là vector adenovirus 26, mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18; và polypeptit kháng nguyên HIV được phân lập có chứa các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7 hoặc các gốc 30-724 của SEQ ID NO: 36. Theo một phương án cụ thể, chế phẩm thứ nhất còn chứa vector adenovirus, tốt hơn là vector adenovirus 26, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 5. Theo phương án cụ thể khác, chế phẩm thứ nhất còn chứa một hoặc nhiều vector adenovirus bổ sung, tốt hơn là vector adenovirus 26, mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 28 và 29.

Khía cạnh chung khác của sáng chế đề cập đến bộ kit có chứa vacxin kết hợp theo phương án của sáng chế.

Các phương án khác của protein vỏ HIV tổng hợp, vector biểu hiện, vector biểu hiện bổ sung, polypeptit kháng nguyên HIV được mã hóa bởi vector biểu hiện, và polypeptit kháng nguyên HIV được phân lập v.v. mà có thể được sử dụng trong vacxin kết hợp theo sáng chế được thảo luận chi tiết ở trên và trong các ví dụ minh họa dưới đây.

Phương Pháp Gây Ra Sự Miễn Dịch Bảo Vệ Chống Lại Sự Nhiễm HIV

Sáng chế cũng đề cập đến phương pháp gây ra đáp ứng miễn dịch chống lại một hoặc nhiều nhánh HIV ở đối tượng cần chủng. Phương pháp được mô tả trong bản mô tả này bao gồm phương pháp môi và thúc đẩy đáp ứng miễn dịch bằng cách sử dụng một hoặc nhiều vector biểu hiện ở dạng kết hợp với một hoặc nhiều polypeptit kháng nguyên được phân lập.

Theo một khía cạnh chung, phương pháp gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng có chứa bước dùng cho đối tượng chế phẩm có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector biểu hiện có chứa axit nucleic mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8. Chế phẩm bất kỳ được mô tả trong bản mô tả này có thể được sử dụng trong phương pháp gây ra đáp ứng miễn dịch chống lại HIV ở đối tượng. Tốt hơn là, chế phẩm có chứa vector adenovirus, tốt hơn là vector adenovirus 26, có chứa axit nucleic mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18. Chế phẩm có thể còn chứa một hoặc nhiều vector bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung và/hoặc một hoặc nhiều polypeptit kháng nguyên HIV được phân lập bổ sung.

Theo một khía cạnh chung khác, phương pháp gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng có chứa bước:

- (i) cho đối tượng sử dụng chế phẩm thứ nhất có chứa lượng hữu hiệu để gây đáp ứng miễn dịch vector biểu hiện mã hoá protein vỏ HIV chứa trình tự axit amin được nêu trong SEQ ID NO: 8 hoặc SEQ ID NO: 8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (a) I529P, (b) K480E, và (c) tổ hợp các đột biến EK479-480RRRR, I529P, A471C và T575C, và chất mang;
- (ii) cho đối tượng sử dụng chế phẩm thứ hai có chứa lượng hữu hiệu để gây đáp ứng miễn dịch polypeptit kháng nguyên HIV được phân lập và chất mang; và
- (iii) tùy ý, cho đối tượng sử dụng lượng hữu hiệu để gây đáp ứng miễn dịch một hoặc nhiều vector biểu hiện bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung,

trong đó các bước (i) và (ii) được thực hiện theo trình tự bất kỳ, với một trong số các bước để mỗi miễn dịch và bước kia để thúc đẩy miễn dịch. Theo phương án của sáng chế, lượng hữu hiệu, tùy chọn của một hoặc nhiều vector biểu hiện bổ sung được dùng cùng với chế phẩm thứ nhất hoặc chế phẩm thứ hai. Theo phương án được ưu tiên, lượng hữu hiệu tùy chọn của một hoặc nhiều vector biểu hiện bổ sung được dùng cùng với chế phẩm thứ nhất.

Theo phương án cụ thể của phương pháp gây ra đáp ứng miễn dịch, chế phẩm thứ nhất có chứa vector adenovirut, tốt hơn là vector adenovirut 26, mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 và vector adenovirut thứ hai, tốt hơn là vector adenovirut 26, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 5; chế phẩm thứ hai có chứa polypeptit kháng nguyên HIV được phân lập có các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO:7 hoặc các gốc 30-724 của SEQ ID NO:36; và một hoặc nhiều vector biểu hiện bổ sung là vector adenovirut, tốt hơn là vector adenovirut 26, mã hoá một hoặc nhiều polypepit kháng nguyên HIV bổ sung có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 28 và 29; trong đó chế phẩm thứ nhất được dùng cho đối tượng, cùng với một hoặc nhiều vector biểu hiện bổ sung, một hoặc nhiều lần mỗi miễn dịch, và chế phẩm thứ hai được dùng cho đối tượng một hoặc nhiều lần để thúc đẩy miễn dịch.

Việc dùng chế phẩm miễn dịch có chứa vector biểu hiện và/hoặc polypeptit kháng nguyên thường là trong cơ, trong da hoặc dưới da. Tuy nhiên, các phương thức sử dụng khác như trong tĩnh mạch, qua trực tràng, da, đường miệng, qua đường mũi, v.v. cũng có thể được dự tính. Việc sử dụng trong cơ của chế phẩm miễn dịch có thể đạt được bằng cách sử dụng kim để tiêm huyền phù của vector biểu hiện, ví dụ như vector adenovirut, và/hoặc polypeptit kháng nguyên. Phương án thay thế là sử dụng dụng cụ tiêm không kim để dùng chế phẩm (bằng cách sử dụng, ví dụ như, Biojector™) hoặc bột đông khô chứa vaccin.

Để tiêm trong cơ, trong tĩnh mạch, da hoặc dưới da, hoặc tiêm ở vị trí đau, vector sẽ ở dạng dung dịch trong nước chấp nhận được ngoài đường tiêu hóa mà không chứa chất gây sốt và có độ pH, tính đẳng trương và độ ổn định thích hợp. Cũng như vậy, polypeptit kháng nguyên được phân lập sẽ ở dạng dung dịch chấp nhận được ngoài đường tiêu hóa có độ pH, tính đẳng trương, và độ ổn định thích hợp. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này có khả năng điều chế dung dịch thích hợp bằng cách sử dụng, ví dụ, tá dược lỏng đẳng trương như Dung Dịch Tiêm Natri Clorua, Dung Dịch Tiêm Ringer, Dung Dịch Tiêm Ringer Lactat. Chất bảo quản, chất ổn định, chất

đệm, chất chống oxy hóa và/hoặc các chất phụ gia khác có thể được bao gồm, khi cần thiết. Chế phẩm giải phóng chậm cũng có thể được sử dụng.

Thông thường, việc dùng chế phẩm vacxin theo phương án của sáng chế là nhằm mục đích phòng ngừa để tạo ra đáp ứng miễn dịch chống lại kháng nguyên HIV trước khi nhiễm hoặc phát triển triệu chứng. Theo các phương án khác, vector biểu hiện, ví dụ như, vector adenovirut, và/hoặc polypeptit kháng nguyên HIV có thể được dùng để phòng ngừa sau phơi nhiễm.

Chế phẩm gây miễn dịch chứa vector biểu hiện, ví dụ như, vector adenovirut, và/hoặc polypeptit kháng nguyên được dùng cho đối tượng, mang lại sự tăng lên cho đáp ứng miễn dịch kháng HIV ở đối tượng. Lượng chế phẩm đủ để gây ra đáp ứng miễn dịch có thể phát hiện được được xác định là “liều lượng hữu hiệu để gây đáp ứng miễn dịch” hoặc “lượng hữu hiệu để gây đáp ứng miễn dịch.” Theo phương án điển hình của sáng chế, đáp ứng miễn dịch là đáp ứng miễn dịch bảo vệ.

Lượng thực sự được dùng, và tỉ lệ và thời gian dùng, sẽ phụ thuộc vào bản chất và mức độ nghiêm trọng của bệnh cần được điều trị. Đơn thuốc điều trị, ví dụ như, quyết định về liều lượng v.v., thuộc về trách nhiệm của bác sĩ đa khoa và bác sĩ y khoa khác, hoặc bác sĩ thú y trong trường hợp thú y, và thường có tính đến rối loạn cần điều trị, tình trạng của bệnh nhân cụ thể, vị trí phân phối, phương pháp sử dụng và các yếu tố khác đã biết đối với bác sĩ. Ví dụ về kỹ thuật và quy trình được đề cập ở trên có thể được tìm thấy trong Remington's Pharmaceutical Sciences, 16th edition, Osol, A. ed., 1980.

Sau khi sản xuất vector adenovirut và chế phẩm tùy chọn của các hạt này thành chế phẩm, vector có thể được dùng cho cá thể, cụ thể là người hoặc động vật linh trưởng khác. Việc phân phối cho động vật có vú không phải là người không cần phải là nhằm mục đích trị, mà có thể là để sử dụng trong trường hợp thí nghiệm, ví dụ trong việc nghiên cứu cơ chế của đáp ứng miễn dịch đối với protein vỏ HIV tổng hợp được biểu hiện bởi vector adenovirut theo sáng chế.

Theo một phương án của phương pháp được bộc lộ, một hoặc nhiều vector adenovirut mã hoá một hoặc nhiều polypeptit kháng nguyên HIV được sử dụng để mỗi đáp ứng miễn dịch. Một hoặc nhiều polypeptit kháng nguyên HIV được phân lập có thể được sử dụng cùng với một hoặc nhiều vector adenovirut cho sự gây miễn dịch mỗi. Sự gây miễn dịch mỗi có thể được dùng chỉ một lần nhưng cũng có thể tùy ý được dùng nhiều lần, ví dụ, việc dùng để mỗi lần đầu ở thời điểm 0, sau đó là việc dùng để mỗi khác khoảng 4-14 tuần, ví dụ như 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 hoặc 14 tuần, hoặc thời điểm bất kỳ giữa đó, sau khi dùng để mỗi lần đầu. Một hoặc nhiều polypeptit kháng nguyên HIV được phân lập tùy ý cùng với một hoặc nhiều adenovirut bổ sung hoặc các

vector khác mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung có thể được sử dụng để thúc đẩy đáp ứng miễn dịch. Sự gây miễn dịch thúc đẩy cũng có thể được dùng một lần hoặc nhiều lần, ví dụ, lần thứ nhất ở khoảng 18- 36, ví dụ như 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 hoặc 36 tuần, hoặc thời điểm bất kỳ giữa đó, sau khi dùng để mỗi lần đầu, sau đó là việc dùng để thúc đẩy khác ở khoảng 36-52, ví dụ như 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59 hoặc 60 tuần, hoặc thời điểm bất kỳ giữa đó, sau khi dùng để mỗi lần đầu. Đáp ứng miễn dịch gây ra bởi sự gây miễn dịch này được theo dõi.

Các phương án của phương pháp được bộc lộ còn đề cập đến chế độ trị liệu miễn dịch thúc đẩy ngắn hơn, có nghĩa là sự gây miễn dịch thúc đẩy cuối cùng được dùng khoảng 22-26 tuần sau khi dùng để mỗi lần đầu. Sự gây miễn dịch mỗi có thể được dùng ở tuần 0. Sự gây miễn dịch thúc đẩy có thể được dùng nhiều lần, ví dụ, lần thứ nhất ở khoảng 7-9 tuần hoặc 11-13 tuần, hoặc ở khoảng 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, hoặc 14 tuần, hoặc thời điểm bất kỳ giữa đó, sau khi dùng để mỗi lần đầu, sau đó là việc dùng để thúc đẩy khác ở khoảng 22-26 tuần, hoặc ở khoảng 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, hoặc 28 tuần, hoặc thời điểm bất kỳ giữa đó, sau khi dùng để mỗi lần đầu. Theo các phương án nhất định, một hoặc nhiều polypeptit kháng nguyên HIV được phân lập được dùng cùng với một hoặc nhiều vector adenovirut để gây đáp ứng miễn dịch mỗi và/hoặc thúc đẩy.

Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này hiểu rõ rằng chế độ trị liệu dùng để mỗi và thúc đẩy có thể được điều chỉnh dựa trên đáp ứng miễn dịch đo được sau khi dùng. Ví dụ, chế phẩm thúc đẩy thường được dùng vài tuần hoặc vài tháng sau khi dùng chế phẩm mỗi, ví dụ, khoảng 2-3 tuần hoặc 4 tuần, hoặc 8 tuần, hoặc 16 tuần, hoặc 20 tuần, hoặc 24 tuần, hoặc 28 tuần, hoặc 30 tuần hoặc 32 tuần hoặc một đến hai năm sau khi dùng chế phẩm mỗi.

Theo phương án của sáng chế, tá dược có thể được dùng cùng với polypeptit kháng nguyên HIV được phân lập dưới dạng một phần của sự gây miễn dịch mỗi và/hoặc thúc đẩy. Tá dược bất kỳ có thể được sử dụng dựa vào sáng chế, và theo các phương án nhất định tá dược là muối nhôm, như AdjuPhos.

Theo phương án được ưu tiên của sáng chế, vector adenovirut được sử dụng trong phương pháp được bộc lộ trong bản mô tả này bao gồm vector rAd26. Tốt hơn là, vector rAd26 mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18 hoặc SEQ ID NO: 19, tốt nhất là SEQ ID NO: 18, được sử dụng để mỗi đáp ứng miễn dịch, một mình hoặc kết hợp với một hoặc nhiều vector rAd26 bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung, như mos1Env có trình tự axit amin được nêu trong SEQ ID NO:5, và polypeptit kháng nguyên HIV được phân lập,

như polypeptit có chứa các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7 hoặc các gốc 30-724 của SEQ ID NO: 36, được sử dụng để thúc đẩy đáp ứng miễn dịch, hoặc ngược lại.

Theo một phương án ví dụ, vector rAd26 mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18 được sử dụng để môi đáp ứng miễn dịch kết hợp với vector rAd26 mã hoá polypeptit kháng nguyên HIV có trình tự axit amin được nêu trong SEQ ID NO: 5. Một hoặc nhiều vector rAd26 bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung có trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 1-4, 28 và 29 cũng có thể được dùng cùng với các vector rAd26 khác để môi đáp ứng miễn dịch. Việc dùng để môi theo các phương án nhất định được dùng hai lần trước khi dùng sự gây miễn dịch thúc đẩy bất kỳ. Polypeptit kháng nguyên HIV được phân lập, chẳng hạn như polypeptit có chứa các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7 (tốt hơn là), hoặc polypeptit có chứa các gốc 30-724 của trình tự axit amin được nêu trong SEQ ID NO: 36, hoặc dạng kết hợp của ít nhất hai polypeptit kháng nguyên HIV này được phân lập, sau đó được dùng để thúc đẩy đáp ứng miễn dịch, và tốt hơn là được dùng nhiều hơn một lần. Tốt hơn là, tá được dùng thêm với polypeptit kháng nguyên HIV được phân lập trong sự gây miễn dịch thúc đẩy.

Theo phương án cụ thể, đáp ứng miễn dịch được môi bằng cách sử dụng bốn kháng nguyên HIV được mã hóa trên vector adenovirut, tốt hơn là vector rAd26, bốn kháng nguyên được mã hóa là: (i) protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18, (ii) polypeptit có trình tự axit amin được nêu trong SEQ ID NO: 5, (iii) polypeptit có trình tự axit amin được nêu trong SEQ ID NO: 28, và (iv) polypeptit có trình tự axit amin được nêu trong SEQ ID NO: 29. Mỗi trong bốn kháng nguyên này có thể được mã hóa trên vector adenovirut riêng rẽ, tốt hơn là vector rAd26, được dùng ở liều lượng tổng cộng bằng khoảng 1, 2, 3, 4, 5, 6, 7, 8, 9, hoặc 10×10^{10} hạt virut (vp), ví dụ như khoảng 5×10^{10} vp (đối với tất cả các vector cùng nhau). Các vector có thể được trộn sơ bộ, ví dụ như ở tỉ lệ 1:1:1:1. Việc dùng để môi có thể được lặp lại sau khi dùng để môi lần đầu, ví dụ như ở 8, 9, 10, 11, 12, 13, 14, 15 hoặc 16 tuần sau khi dùng để môi lần đầu. Theo phương án này, đáp ứng miễn dịch được thúc đẩy bằng cách sử dụng cùng vacxin vector adenovirut dùng để môi cùng với protein gp140 Env HIV được phân lập, ví dụ như protein gp140 nhánh C (có chứa các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7), hoặc protein gp140 khảm (có chứa các gốc 30-724 của trình tự axit amin được nêu trong SEQ ID NO: 36), hoặc protein gp140 nhánh C và protein gp140 khảm, ở liều lượng tổng cộng bằng khoảng 50-300 μ g protein, ví dụ như 50, 100, 150, 200, 250, hoặc 300 microgam, hoặc lượng bất

kỳ ở giữa chúng, của protein gp140 khảm, hoặc ví dụ như 50, 100, 150, 200, 250, hoặc 300 microgam, hoặc lượng bất kỳ ở giữa, của dạng kết hợp protein gp140 nhánh C và protein gp140 khảm (ví dụ như ở tỉ lệ 1:1, được trộn với nhau hoặc được dùng riêng rẽ). Tốt hơn là protein gp140 được dùng cùng với tá dược, ví dụ như nhôm phosphat. Việc sử dụng adenovirut cộng với protein gp140 để thúc đẩy đáp ứng miễn dịch có thể được thực hiện ở khoảng 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 hoặc 30 tuần, hoặc ở thời điểm bất kỳ giữa đó, sau khi dùng để mỗi lần đầu. Việc dùng để thúc đẩy có thể được lặp lại, ví dụ như ở khoảng 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53 hoặc 54 tuần, hoặc thời điểm bất kỳ giữa đó, sau khi dùng để mỗi lần đầu. Tất cả các lần dùng theo phương án này tốt hơn là được thực hiện qua đường trong cơ.

Phương án thực hiện sáng chế

Phương án 1 là axit nucleic mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8, hoặc SEQ ID NO: 8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (i) I529P, (ii) K480E, và (iii) tổ hợp các đột biến EK479-480RRRR, I529P, A471C và T575C.

Phương án 2 là axit nucleic theo phương án 1, trong đó protein vỏ HIV tổng hợp còn chứa trình tự tín hiệu, ví dụ trình tự tín hiệu có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO:9 đến 12, tốt hơn là SEQ ID NO: 9.

Phương án 3 là axit nucleic theo phương án 1 hoặc 2, trong đó protein vỏ HIV tổng hợp còn chứa vùng xuyên màng, ví dụ vùng xuyên màng có trình tự axit amin được nêu trong SEQ ID NO: 13, tốt hơn là protein vỏ HIV tổng hợp còn chứa SEQ ID NO:37 được dung hợp với đầu tận cùng C của SEQ ID NO:8 và đầu tận cùng N của vùng xuyên màng.

Phương án 4 là axit nucleic theo phương án 3, trong đó protein vỏ HIV tổng hợp còn chứa mảnh của vùng tế bào chất, tốt hơn là mảnh của vùng tế bào chất có chứa trình tự axit amin được nêu trong SEQ ID NO: 14, hoặc các gốc axit amin 1-4 của chúng (tức là, NRVR).

Phương án 5 là axit nucleic theo phương án bất kỳ trong số các phương án từ 1 đến 4 nêu trên, trong đó protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18.

Phương án 6 là axit nucleic theo phương án 1 hoặc 2, trong đó protein vỏ HIV tổng hợp (a) còn chứa vùng trime hoá được chọn từ nhóm gồm có GCN4, fibritin (miền foldon), ví dụ vùng trime hoá có trình tự axit amin được nêu trong SEQ ID NO: 15 hoặc SEQ ID NO:16, tốt hơn là SEQ ID NO: 15, hoặc (b) có chứa SEQ ID NO:8 với tổ hợp các đột biến: EK479-480RRRR, I529P, A471C và T575C.

Phương án 7 là axit nucleic theo phương án 6, trong đó protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 19.

Phương án 8 là axit nucleic theo phương án 5, trong đó protein vỏ HIV tổng hợp gồm có trình tự axit amin được nêu trong SEQ ID NO: 18.

Phương án 9 là axit nucleic theo phương án 7, trong đó protein vỏ HIV tổng hợp gồm có trình tự axit amin được nêu trong SEQ ID NO: 19.

Phương án 10 là vectơ có chứa axit nucleic theo phương án bất kỳ trong số các phương án từ 1 đến 9, trong đó axit nucleic được liên kết có điều khiển với trình tự khởi động.

Phương án 11 là vectơ theo phương án 10 là vectơ virus, tốt hơn là vectơ adenovirus, và tốt hơn nữa là vectơ adenovirus 26.

Phương án 12 là tế bào được phân lập có chứa vectơ theo phương án 10 hoặc phương án 11.

Phương án 13 là chế phẩm có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vectơ theo phương án 10 hoặc điểm 11, và chất mang.

Phương án 14 là vaccin kết hợp, có chứa chế phẩm thứ nhất có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vectơ adenovirus, tốt hơn là vectơ adenovirus 26, mã hoá protein vỏ HIV tổng hợp có trình tự axit amin được nêu trong SEQ ID NO: 18, chế phẩm thứ hai có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vectơ adenovirus thứ hai, tốt hơn là vectơ adenovirus 26 thứ hai, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 5, và tùy ý ít nhất một chế phẩm bổ sung có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của ít nhất một vectơ được chọn từ nhóm gồm có vectơ mã hoá polypeptit kháng nguyên có trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 1-4, 28 và 29, và polypeptit kháng nguyên HIV được phân lập có chứa các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7, hoặc các gốc 30-724 của trình tự axit amin được nêu trong SEQ ID NO: 36, trong đó chế phẩm thứ nhất, chế phẩm thứ hai và chế phẩm bổ sung có mặt trong cùng chế phẩm hoặc trong một hoặc nhiều chế phẩm khác nhau.

Phương án 15 là phương pháp gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng cần chúng, phương pháp này có chứa bước dùng cho đối tượng chế phẩm theo phương án 13 hoặc vaccin kết hợp theo phương án 14.

Phương án 16 là chế phẩm theo phương án 13 hoặc vaccin kết hợp theo phương án 14, có chứa vectơ adenovirus, tốt hơn là vectơ adenovirus 26, mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18, vectơ adenovirus thứ hai, tốt hơn là vectơ adenovirus 26, mã hoá polypeptit kháng nguyên HIV có chứa

trình tự axit amin được nêu trong SEQ ID NO: 5, một hoặc nhiều vector adenovirut bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên bổ sung có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 1-4, 28 và 29, và polypeptit kháng nguyên HIV được phân lập có chứa các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7 hoặc các gốc 30-724 của SEQ ID NO: 36, để sử dụng trong việc gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV).

Phương án 17 là protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 hoặc SEQ ID NO: 8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (i) I529P, (ii) K480E, và (iii) tổ hợp các đột biến EK479-480RRRR, I529P, A471C và T575C.

Phương án 18 là protein vỏ HIV tổng hợp theo phương án 17, có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 với tổ hợp các đột biến EK479-480RRRR, I529P, A471C và T575C, hoặc các gốc 30-704 của trình tự axit amin trong SEQ ID NO: 18 hoặc các gốc 30-686 của SEQ ID NO: 19.

Phương án 19 là chế phẩm theo phương án 13, còn chứa một hoặc nhiều vector biểu hiện bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung, và/hoặc một hoặc nhiều polypeptit kháng nguyên HIV được phân lập.

Phương án 20 là chế phẩm theo phương án 13 có chứa vector adenovirut, tốt hơn là vector adenovirut 26, mã hoá protein vỏ HIV tổng hợp gồm có trình tự axit amin được nêu trong SEQ ID NO: 18.

Phương án 21 là chế phẩm theo phương án 20 còn chứa vector adenovirut thứ hai, tốt hơn là vector adenovirut 26, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 5, và tùy ý một hoặc nhiều vector adenovirut bổ sung, tốt hơn là vector adenovirut 26, mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung có chứa trình tự axit amin của các SEQ ID NO: 1-4, 28 và 29.

Phương án 22 là phương pháp tạo ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng cần chủng, phương pháp này có chứa bước dùng cho đối tượng chế phẩm theo phương án bất kỳ trong số các phương án 19, 20, hoặc 21.

Phương án 23 là phương pháp sản xuất chế phẩm hoặc vacxin kết hợp, có chứa bước kết hợp vector theo phương án 10 hoặc phương án 11 với chất mang, và tùy ý một hoặc nhiều vector bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung và/hoặc một hoặc nhiều polypeptit kháng nguyên HIV được phân lập trong một hoặc nhiều chế phẩm, cùng với chất mang.

Phương án 24 là vacxin kết hợp để gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng, có chứa:

(i) chế phẩm thứ nhất có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector biểu hiện mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 hoặc SEQ ID NO: 8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (i) I529P, (ii) K480E, và (iii) tổ hợp các đột biến EK479-480RRRR, I529P, A471C và T575C, và chất mang; và

(ii) chế phẩm thứ hai có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của polypeptit kháng nguyên HIV được phân lập và chất mang,

trong đó một trong số chế phẩm thứ nhất và chế phẩm thứ hai là để miễn dịch và chế phẩm kia là để thúc đẩy sự miễn dịch, và

trong đó vaccin kết hợp tùy ý còn chứa lượng hữu hiệu để gây đáp ứng miễn dịch của một hoặc nhiều vector biểu hiện bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung, và một hoặc nhiều vector biểu hiện bổ sung này được bao gồm trong chế phẩm thứ nhất hoặc chế phẩm thứ hai hoặc một hoặc nhiều chế phẩm bổ sung được sử dụng cùng với chế phẩm thứ nhất hoặc chế phẩm thứ hai.

Phương án 25 là vaccin kết hợp theo phương án 24, trong đó chế phẩm thứ nhất có chứa vector adenovirut, tốt hơn là vector adenovirut 26, mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18; polypeptit kháng nguyên HIV được phân lập có chứa các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7, hoặc các gốc 30-724 của SEQ ID NO: 36; và một hoặc nhiều vector biểu hiện bổ sung là vector adenovirut, tốt hơn là vector adenovirut 26, mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 1-5, 28 và 29.

Phương án 26 là phương pháp gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng cần chúng, phương pháp này có chứa bước:

(i) cho đối tượng sử dụng chế phẩm thứ nhất có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector biểu hiện mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 hoặc SEQ ID NO: 8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (i) I529P, (ii) K480E, và (iii) tổ hợp các đột biến EK479-480RRRR, I529P, A471C và T575C, và chất mang;

(ii) cho đối tượng sử dụng chế phẩm thứ hai có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của polypeptit kháng nguyên HIV được phân lập và chất mang; và

(iii) tùy ý, cho đối tượng sử dụng lượng hữu hiệu để gây đáp ứng miễn dịch của một hoặc nhiều vector biểu hiện bổ sung mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung,

trong đó các bước (i) và (ii) được thực hiện theo trình tự bất kỳ, với một trong số các bước là để miễn dịch và bước kia để thúc đẩy đáp ứng miễn dịch, và

tốt hơn là, lượng hữu hiệu, tùy chọn của một hoặc nhiều vector biểu hiện bổ sung được dùng cùng với chế phẩm thứ nhất.

Phương án 27 là phương pháp theo phương án 26, trong đó chế phẩm thứ nhất có chứa vector adenovirut, tốt hơn là vector adenovirut 26, mã hoá protein vỏ HIV tổng hợp có trình tự axit amin được nêu trong SEQ ID NO: 18 và vector adenovirut thứ hai, tốt hơn là vector adenovirut 26, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 5; chế phẩm thứ hai có chứa polypeptit kháng nguyên HIV được phân lập có các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7, hoặc các gốc 30-724 của SEQ ID NO: 36; và một hoặc nhiều vector biểu hiện bổ sung tùy chọn là vector adenovirut, tốt hơn là vector adenovirut 26, mã hoá một hoặc nhiều polypeptit kháng nguyên HIV bổ sung có chứa trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 28 và 29; trong đó chế phẩm thứ nhất được cho đối tượng sử dụng, tùy ý cùng với một hoặc nhiều vector biểu hiện bổ sung, một hoặc nhiều lần mỗi miễn dịch, và chế phẩm thứ hai được cho đối tượng sử dụng một hoặc nhiều lần để thúc đẩy miễn dịch.

Phương án 28 là protein vỏ HIV tổng hợp gồm có trình tự axit amin được nêu trong SEQ ID NO: 18 hoặc SEQ ID NO: 19, có hoặc không có trình tự tín hiệu.

Phương án 29 là vacxin kết hợp có chứa một hoặc nhiều vector cùng với nhau có chứa trình tự axit nucleic mã hoá (i) protein vỏ HIV tổng hợp thứ nhất có chứa trình tự axit amin được nêu trong SEQ ID NO: 8 và (ii) protein vỏ HIV thứ hai có chứa trình tự axit amin được nêu trong SEQ ID NO: 5.

Phương án 30 là vacxin kết hợp theo phương án 29, trong đó protein vỏ HIV tổng hợp thứ nhất có chứa trình tự axit amin được nêu trong SEQ ID NO: 18.

Phương án 31 là vacxin kết hợp có chứa các thành phần sau đây:

(i) vector Ad26 mã hoá protein vỏ HIV tổng hợp gồm có trình tự axit amin được nêu trong SEQ ID NO: 18; và

(ii) vector Ad26 mã hoá protein vỏ HIV gồm có trình tự axit amin được nêu trong SEQ ID NO: 5.

Phương án 32 là vacxin kết hợp theo phương án 31, còn chứa các thành phần sau đây:

(iii) vector Ad26 mã hoá kháng nguyên HIV gồm có trình tự axit amin được nêu trong SEQ ID NO: 28.

Phương án 33 là vacxin kết hợp theo phương án 31 hoặc 32, còn chứa các thành phần sau đây:

(iv) vector Ad26 mã hoá kháng nguyên HIV gồm có trình tự axit amin được nêu trong SEQ ID NO: 29.

Phương án 34 là vaccin kết hợp theo phương án bất kỳ trong số các phương án từ 31 đến 33, còn chứa các thành phần sau đây:

(v) polypeptit kháng nguyên HIV được phân lập có các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7, hoặc các gốc 30-724 của trình tự axit amin được nêu trong SEQ ID NO: 36, tùy ý còn chứa tá được.

Phương án 35 là phương pháp gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng người cần chủng, phương pháp này có chứa bước:

(a) cho đối tượng sử dụng: (i) vectơ rAd26 mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18; (ii) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 5; (iii) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 28; và (iv) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 29; tốt hơn là trong đó các vectơ rAd26 được dùng ở tỉ lệ khoảng 1:1:1:1 ở liều lượng tổng cộng bằng khoảng $1-10 \times 10^{10}$ hạt virus (vp), ví dụ như 5×10^{10} vp;

(b) lặp lại bước (a) ở khoảng 10-14 tuần, ví dụ như ở 12 tuần sau bước (a);

(c) cho đối tượng sử dụng: (i) vectơ rAd26 mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18; (ii) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 5; (iii) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 28; (iv) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 29; (v) protein gp140 HIV được phân lập có trình tự của các axit amin 30-708 của SEQ ID NO: 7; và (vi) tá được nhôm phosphat; tốt hơn là trong đó các vectơ rAd26 được dùng ở tỉ lệ khoảng 1:1:1:1 ở liều lượng tổng cộng bằng khoảng $1-10 \times 10^{10}$ hạt virus (vp), ví dụ như 5×10^{10} vp và trong đó protein gp140 HIV được phân lập được dùng ở liều lượng bằng khoảng 50-300 microgam, ví dụ như 250 microgam; ở khoảng 20-28 tuần, ví dụ như ở 24 tuần sau bước (a); và

(d) lặp lại bước (c) ở khoảng 42-54 tuần, ví dụ như ở 48 tuần sau bước (a);

Phương án 36 là phương pháp gây ra đáp ứng miễn dịch chống lại virus gây suy giảm miễn dịch ở người (HIV) ở đối tượng người cần chủng, phương pháp này có chứa bước:

(a) cho đối tượng sử dụng: (i) vectơ rAd26 mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18; (ii) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 5; (iii) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 28; và (iv) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID

NO: 29; tốt hơn là trong đó các vectơ rAd26 được dùng ở tỉ lệ khoảng 1:1:1:1 ở liều lượng tổng cộng bằng khoảng $1-10 \times 10^{10}$ hạt virus (vp), ví dụ như 5×10^{10} vp;

(b) lặp lại bước (a) ở khoảng 10-14 tuần, ví dụ như ở 12 tuần sau bước (a);

(c) cho đối tượng sử dụng: (i) vectơ rAd26 mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18; (ii) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 5; (iii) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 28; (iv) vectơ rAd26 mã hoá kháng nguyên có chứa trình tự axit amin được nêu trong SEQ ID NO: 29; (v) protein gp140 HIV được phân lập có trình tự của các axit amin 30-708 của SEQ ID NO: 7; (vi) protein gp140 HIV được phân lập có trình tự của các axit amin 30-724 của SEQ ID NO: 36; và (vii) tá được nhôm phosphat; tốt hơn là trong đó các vectơ rAd26 được dùng ở tỉ lệ khoảng 1:1:1:1 ở liều lượng tổng cộng bằng khoảng $1-10 \times 10^{10}$ hạt virus (vp), ví dụ như 5×10^{10} vp và trong đó protein gp140 HIV được phân lập được dùng ở tỉ lệ bằng khoảng 1:1 ở liều lượng tổng cộng bằng khoảng 50-300 microgam, ví dụ như 250 microgam; ở khoảng 20-28 tuần, ví dụ như ở 24 tuần sau bước (a); và

(d) lặp lại bước (c) ở khoảng 42-54 tuần, ví dụ như ở 48 tuần sau bước (a);

Ví dụ thực hiện sáng chế

Ví dụ 1: Thiết kế của trình tự kháng nguyên vỏ HIV

Một vài trình tự kháng nguyên vỏ HIV được thiết kế có độ tương tự trình tự so với kháng nguyên HIV khảm mos2Env (SEQ ID NO: 6; cũng được mô tả trước đây trong WO 2010/059732). Các trình tự liên kết màng, được thiết kế mới, được dựa trên (sự kết hợp của) các trình tự kiểu đại tự nhiên đầy đủ từ protein vỏ HIV, hoặc thể khảm của trình tự mos2Env và trình tự protein vỏ HIV kiểu đại. Ngoài trình tự protein vỏ chiều dài đầy đủ (xem FIG. 1A), các trình tự có phần cắt đầu tận cùng C của vùng tế bào chất cũng được thiết kế (xem, ví dụ như, FIG. 1C). Cũng xem ví dụ như, Schiernle et al., PNAS 1997; Abrahamyan et al., J Virol 2005; Edwards et al., *J. Virology*, 2002, 76:2683-2691. Các biến thể hòa tan cũng được điều chế bằng phần cắt đầu tận cùng C trước vùng xuyên màng (TM), được thay thế bằng vùng trime hoá, như vùng trime hoá GCN4 (xem, ví dụ như, FIG. 1B). Các biến thể hòa tan này được biến đổi thêm thành biến thể chuỗi đơn lẻ bằng sự đột biến của vị trí phân cắt furin, do đó ức chế việc xử lý của vùng ngoại bào của protein vỏ thành các đơn vị con gp120 và gp41.

Trong số tất cả các cấu trúc được tạo ra và được thử nghiệm, các cấu trúc dựa trên C4 có tính chất tối ưu nhất, ví dụ như, khả năng sản xuất, sự gấp nếp, khả năng gây miễn dịch, v.v. tốt và chúng được chọn cho các nghiên cứu tiếp theo. Biến thể hòa tan của cấu trúc C4 có vùng trime hoá GCN4 thay cho vùng xuyên màng (sC4, FIG. 1B), và biến thể có chứa mảnh có 7 axit amin của vùng tế bào chất (C4D7, FIG. 1C) cũng

được tạo ra và được thử nghiệm trong các nghiên cứu tiếp theo. Trình tự axit amin của C4, sC4, và C4D7 lần lượt được thể hiện trong các SEQ ID NO: 17, 19, và 18. Các trình tự mã hoá các protein này lần lượt được thể hiện trong các SEQ ID NO: 25, 27, và 26. Cấu trúc C1 có trình tự vùng ngoại bào dựa trên trình tự mos2Env (SEQ ID NO: 6). Biến thể hòa tan của cấu trúc C1 có vùng trime hoá GCN4 thay cho vùng xuyên màng (sC1), và biến thể có chứa mảnh có 7 axit amin của vùng tế bào chất (C1D7), tương tự với sC4 và C4D7 như lần lượt được thể hiện trên các FIG. 1B và 1C, cũng được tạo ra. Cấu trúc C1 và biến thể của nó được sử dụng trong các nghiên cứu tiếp theo nhằm mục đích so sánh, vì chúng về cơ bản dựa trên trình tự mos2Env trong tình trạng kỹ thuật. Trình tự axit amin của C1, sC1 và C1D7 lần lượt được thể hiện trong các SEQ ID NO: 31, 30, và 32. Trình tự axit nucleic mã hoá các protein này lần lượt được thể hiện trong các SEQ ID NO: 34, 33, và 35. Các cấu trúc khác được thử nghiệm ít tối ưu hơn so với các cấu trúc dựa trên cấu trúc C4, và không được phát triển tiếp.

Ví dụ 2: Sự Biểu Hiện Và Gấp nếp Của Protein Vỏ HIV Tổng Hợp

Mức độ biểu hiện, gấp nếp, và sự biểu hiện bề mặt tế bào của protein vỏ HIV tổng hợp được đo.

Mức Độ Biểu Hiện

Tế bào HEK293F được chuyển nạp tạm thời bằng plasmit mã hoá các protein vỏ HIV tổng hợp hòa tan sC1 và sC4 như được mô tả trong Ví dụ 1. Mức độ biểu hiện của protein hòa tan được đo trong dịch nổi bề mặt bằng cách sử dụng thẩm tách Western định lượng (QWB). Kết quả được thể hiện trên FIG. 2. Mức độ biểu hiện thấp đối với sC1 (mà về cơ bản tương ứng với mos2Env có vùng xuyên màng được bổ sung) phù hợp với nhận thức gần đây của chúng tôi về mos2Env. Như đã được chứng minh bởi kết quả này, biến thể sC4 theo sáng chế thể hiện mức độ biểu hiện cao hơn đáng kể so với biến thể sC1 (đối chứng).

Sự Gấp nếp Protein

Sự gấp nếp protein được thử nghiệm bằng cách đo sự liên kết của protein vỏ HIV tổng hợp hòa tan với kháng thể (MAb 17b) đã biết là liên kết với vị trí liên kết đồng thụ thể của protein vỏ HIV, mà chỉ được bộc lộ ra sau sự liên kết của CD4, bằng thử nghiệm hấp thụ miễn dịch liên kết enzym (ELISA). Cụ thể là, liên kết của sC4 được tinh chế được thử nghiệm về sự liên kết với MAb 17b có sự liên kết trước của sC4 với CD4, và không có sự liên kết trước của sC4 với CD4. sC1 được tinh chế được sử dụng làm đối chứng. Sự liên kết của MAb 17b với sC4 mà không có sự liên kết CD4 trước với protein vỏ là dấu hiệu của protein vỏ không gấp nếp một phần hoặc được khởi động trước (tức là, Env không ổn định mà tuân theo hình dạng “mở” khi không có sự liên kết CD4). Kết quả của thử nghiệm ELISA được thể hiện trên FIG. 3A và FIG. 3B.

Như thể hiện trên FIG. 3B, sC4 thể hiện sự liên kết mạnh với MAb 17b có sự liên kết trước với CD4, nhưng không có sự liên kết có thể phát hiện được với MAb 17b mà không có sự liên kết trước với CD4. Ngược lại, như được thể hiện trên FIG. 3A, sC1 thể hiện sự liên kết yếu hơn nhiều với MAb 17b cả có và không có sự liên kết trước với CD4. Kết quả này gợi ý rằng sC4 có mô hình gấp nếp đúng, không có sự bộc lộ của vị trí liên kết đồng thụ thể trước khi liên kết CD4.

Sự gấp nếp protein cũng được phân tích bằng điện di gel polyacrylamit tự nhiên (PAGE) của sC1 và sC4 để đánh giá cấu trúc bậc bốn của các biến thể protein hòa tan, và sự tạo thành cầu disulphua không chính xác có thể có giữa các protome. Sau khi điện di trên gel tự nhiên, protein trong gel được phát hiện bằng phân tích thẩm tách Western. Như được thể hiện bởi kết quả trên FIG. 4, phần lớn sC4 có mặt ở trạng thái trime, mà là cấu trúc bậc bốn đúng.

Cùng với nhau, các kết quả của thí nghiệm gấp nếp protein chứng tỏ rằng protein vỏ HIV tổng hợp hòa tan sC4 có biên dạng gấp nếp mong muốn, được cải thiện khi so với biên dạng gấp nếp của kháng nguyên mos2Env hiện nay (được thể hiện bởi sC1).

Sự biểu hiện bề mặt tế bào

Sự biểu hiện bề mặt tế bào của các biến thể liên kết màng của protein vỏ HIV C1 (chiều dài đầy đủ), C4 (chiều dài đầy đủ, xem FIG. 1A), C1D7, và C4D7 cũng được nghiên cứu. Tế bào HEK293T được chuyển nạp tạm thời chỉ với plasmid mã hoá eGFP (đối chứng âm, NC), hoặc với plasmid mã hoá eGFP cùng với cấu trúc biểu hiện mã hoá biến thể protein vỏ HIV. Hai ngày sau chuyển nạp, tế bào được đưa đi phân tích phân loại tế bào được hoạt hóa huỳnh quang (FACS) khi bộc lộ với một vài kháng thể đa dòng và kháng thể đơn dòng định hướng chống lại gp120, và kháng thể thứ cấp, và sau đó thử nghiệm mức độ biểu hiện bề mặt tế bào của protein vỏ. Chất lượng của biến thể vỏ được đánh giá bằng cách xác định mức độ biểu hiện nói chung bằng cách sử dụng kháng thể đa dòng kháng gp120, và bằng cách đánh giá sự liên kết tương đối của các kháng thể trung hòa rộng PG9 và PG16, mà phụ thuộc cấu trúc bậc bốn, và liên kết ưu tiên với trime vỏ được gấp nếp đúng.

Kết quả của thí nghiệm biểu hiện bề mặt tế bào được thể hiện trên FIG. 5. Mức độ biểu hiện bề mặt của các biến thể bị cắt cụt C1D7 và C4D7 như đo được bằng cách sử dụng kháng thể kháng gp120, là cao hơn nhiều so với mức độ biểu hiện bề mặt của các đối tác chiều dài đầy đủ của chúng, lần lượt là C1 và C4. Điều này xác nhận rằng sự làm khuyết 144 gốc từ đầu tận cùng carboxy của Env làm tăng mức độ biểu hiện bề mặt của vỏ. Cấu trúc C4 chiều dài đầy đủ theo sáng chế cũng thể hiện sự liên kết PG9 và PG16 được cải thiện khi so với C1 chiều dài đầy đủ, gợi ý rằng trình tự vỏ C4 được gấp nếp thích hợp (tức là, trime) trên bề mặt tế bào.

Kết quả này cũng chứng tỏ rằng biến thể C1D7, mà về cơ bản là Mos2Env với vùng xuyên màng được bổ sung và 7 axit amin của vùng tế bào chất, có thể được biểu hiện bề mặt trên tế bào HEK293T. Điều này trái ngược với cấu trúc hòa tan ở Ad26.mos2Env, mà không thể được biểu hiện ở mức độ có thể phát hiện được trên bề mặt khi được chuyển nạp vào tế bào A549. Tuy nhiên, sự liên kết tương đối với PG9 và PG16 vừa đủ để có thể phát hiện được ở trên mức nền, gợi ý rằng trình tự vỏ C1D7 được gấp nếp kém và có khả năng là không thể hiện dưới dạng trimer nguyên vẹn trên bề mặt tế bào.

Nhìn chung, biến thể vỏ C4D7 có biên dạng liên kết kháng thể tối ưu nhất, với sự biểu hiện gp120 cao hơn so với đối tác C4 chiều dài đầy đủ của nó, và với sự liên kết PG9 và PG16 tăng lên hơn 15 lần khi so với C1 và C1D7 (FIG. 5).

Ví dụ 3: Độ ổn định của vector mã hoá trình tự vỏ HIV

Công việc trước đó ở phòng thí nghiệm của chúng tôi (không công bố) chỉ ra rằng vector adenovirus 26 (Ad26) mã hoá trình tự kháng nguyên mos2Env cho thấy có tỉ lệ VP/IU tương đối cao (chỉ ra chất lượng thấp của mẻ sản phẩm adenovirus) và hơn nữa các vector này thể hiện vấn đề về độ ổn định. Theo đó, điều quan trọng là phải thử nghiệm độ ổn định của các cấu trúc protein vỏ HIV tổng hợp theo sáng chế trên nền adenovirus.

Vector Ad26 tái tổ hợp (rAd26) mã hoá trình tự kháng nguyên HIV theo sáng chế C4, C4D7, và sC4 như được mô tả ở trên trong Ví dụ 1 được tạo ra trong tế bào PER.C6 (lần lượt được đề cập đến dưới dạng “rAd26.C4”, “rAd26.C4D7”, và “rAd26.sC4”). Dòng (mảng) vector được chọn lựa và mở rộng quy mô để tạo ra các mẻ nghiên cứu. Tối đa là 5 dòng (mảng) virus được mở rộng quy mô đến định dạng T25 và được cấy truyền theo dây cho 10 lần cấy truyền ở định dạng T25 (các lần cấy truyền 1-3 là các bước chuyển nạp và tinh chế mảng, sau đó là 10 lần cấy truyền ở định dạng T25, dẫn đến tổng cộng 13 lần cấy truyền). Độ ổn định di truyền được đánh giá ở lần cấy truyền virus số (vpn) 3, 5, 10 và 13 bằng thử nghiệm PCR bằng gen chuyển E1, sau đó là xác định trình tự ở vpn 13. Kết quả được thể hiện trên FIG. 6.

Vector rAd26 mã hoá C4 (rAd26.C4) chiều dài đầy đủ cho thấy các đặc điểm phát triển kém, như xác định được bằng việc không có tác dụng gây bệnh tế bào (CPE) đầy đủ trong 2-3 ngày; sự không ổn định di truyền, như được xác định bằng sự khuyết vùng bằng gen chuyển E1; hoặc dạng kết hợp của chúng (FIG. 6). Do đặc điểm phát triển kém và sự không ổn định di truyền quan sát được, vector mã hoá C4 chiều dài đầy đủ này không được theo đuổi thêm.

Ngược lại, đối với vector rAd26 mã hoá C4D7 (rAd26.C4D7) và sC4 (rAd26.sC4), tất cả các mảng được nhân giống vẫn ổn định về mặt di truyền trong quá

trình thí nghiệm (FIG. 6). Do đó, các cấu trúc sC4 và C4D7 mới tốt hơn so với cấu trúc mos2Env ban đầu về độ ổn định trong nền vectơ adenovirut. Thử nghiệm độ ổn định di truyền lên đến vpn 13 biểu thị sự nhân giống vượt xa một vài lần cấy truyền so với được sử dụng trong sự điều chế quy mô công nghiệp của vectơ.

Ví dụ 4: Sự biểu hiện và tính kháng nguyên *in vivo* của trình tự vỏ HIV trong vectơ adenovirut

Sự biểu hiện và tính kháng nguyên của rAd26.C4D7 và rAd26.sC4 được đánh giá riêng rẽ hoặc kết hợp với vectơ Ad26 tái tổ hợp mã hoá mos1Env (SEQ ID NO: 5) (sau đây gọi là “rAd26.mos1Env”) trong tế bào A549 được tải nạp vectơ (dòng tế bào người) *in vitro* (dữ liệu không được thể hiện). Phân tích đếm tế bào chứng minh rằng tất cả các kháng nguyên được biểu hiện trong môi trường nuôi cấy tế bào được tải nạp với 2×10^4 hạt virut (vp) của kháng nguyên vỏ đơn lẻ làm đối chứng, hoặc với 1×10^4 vp của 2 kháng nguyên Env kết hợp bằng cách tải nạp adenovirut. Tất cả các tải nạp chứa thêm liều đơn (1×10^4 vp) của vectơ adenovirut mã hoá mos1GagPol (“rAd26.mos1GagPol”) và mos2GagPol (“rAd26.mos2GagPol”) (Barouch et al, *Nat Med* 2010, 16:319-323), sao cho vectơ kết hợp được đánh giá thể hiện cùng tỉ lệ tương đối của các vectơ adenovirut khác nhau khi dự định để sử dụng tiền lâm sàng và lâm sàng. Tốt hơn là, vectơ mã hoá protein vỏ HIV tổng hợp theo sáng chế được kết hợp với vectơ mã hoá kháng nguyên mos1GagPol và kháng nguyên mos2GagPol để sử dụng lâm sàng.

Sự kết hợp của rAd26.mos1Env và rAd26.C4D7 đạt được mức độ phủ tối đa của epitop được đánh giá như được xác định bằng sự liên kết kháng thể đơn dòng. Cụ thể là, sự bộc lộ của epitop PG16, được góp phần bởi sự biến đổi với Ad26.C4D7 hứa hẹn đối với việc sử dụng vaccin vì PG16 biểu thị kháng thể đơn dòng trung hòa rộng nhận diện vùng vòng V1/V2 của Env HIV-1 (Walker et al, *Science*. 2009). Do đó, protein vỏ HIV tổng hợp theo sáng chế có nguồn gốc từ trình tự C4 làm tăng độ rộng của đáp ứng miễn dịch chống lại protein vỏ HIV khi so với đáp ứng miễn dịch được tạo ra chỉ bởi mos1Env. Kháng thể được kích thích bởi bắc xin định hướng về phía vùng protein vỏ đã được thể hiện là tương quan với sự bảo vệ khỏi sự nhiễm HIV-1 trong nghiên cứu RV144 (Haynes et al, *N Engl J Med*. 2012), và do đó protein vỏ HIV tổng hợp theo sáng chế là ứng viên hứa hẹn để bao gồm trong chế độ trị liệu vaccin HIV.

Ví dụ 5: Khả năng gây miễn dịch của vectơ mã hoá protein vỏ HIV tổng hợp

Trình tự protein vỏ HIV tổng hợp theo sáng chế trong nền vectơ Ad26 được thử nghiệm ở thỏ để xác định xem liệu các cấu trúc này có phải là phương án gây miễn dịch thay thế đối với cấu trúc rAd26.mos2Env hay không.

Khả năng gây miễn dịch của vectơ adenovirut mã hoá mos1Env (rAd26.mos1Env; SEQ ID NO: 5) được thử nghiệm một mình, và kết hợp với vectơ adenovirut mã hoá protein vỏ HIV tổng hợp theo sáng chế (rAd26.C4D7 và rAd26.sC4; có chứa SEQ ID NO: 8, cụ thể lần lượt là các SEQ ID NO: 18 và 19). Trong tất cả các trường hợp, vectơ adenovirut 26 mã hoá kháng nguyên mos1GagPol và kháng nguyên mos2GagPol (lần lượt là rAd26.mos1GagPol [SEQ ID NO: 28] và rAd26.mos2GagPol [SEQ ID NO: 29]) cũng được dùng. Cụ thể hơn là, khả năng gây miễn dịch của một mình rAd26.mos1Env (vaccine ba thành phần: rAd26.mos1GagPol, rAd26.mos2GagPol và rAd26.mos1Env) được so với khả năng gây miễn dịch của rAd26.mos1Env ở dạng kết hợp với một trong số rAd26.C4D7 hoặc rAd26.sC4 (vaccine bốn thành phần: dùng rAd26.mos1GagPol, rAd26.mos2GagPol, rAd26.mos1Env và rAd26.C4D7; hoặc dùng rAd26.mos1GagPol, rAd26.mos2GagPol, rAd26.mos1Env và rAd26.sC4). Việc so sánh này của vaccine ba thành phần, mà thiếu vectơ bất kỳ mã hoá protein vỏ HIV tổng hợp theo sáng chế, với vaccine bốn thành phần, mà chứa vectơ mã hoá protein vỏ HIV tổng hợp theo sáng chế, cho phép xác định xem liệu protein vỏ HIV của sáng chế có góp phần vào độ rộng của sự bảo vệ hay không.

Việc dùng được thực hiện trong chế độ trị liệu vaccine, trong đó các vectơ Ad26 này được dùng ở các tuần 0 và 6 làm mỗi kíp, và protein gp140 nhánh C (protein Env gp140 ba thành phần có SEQ ID NO: 7 mà không có trình tự peptit tín hiệu gồm các gốc 1-29, cũng xem WO 2010/042942) vào các tuần 12 và 18 làm sự thúc đẩy kíp (xem tài liệu ví dụ như Barouch et al, 2015, Science 349: 320-324). Bảng 1 mô tả chế độ trị liệu vaccine được sử dụng cho nghiên cứu này. rAd26.Empty dùng để chỉ vectơ đối chứng thiếu gen bất kỳ mã hoá trình tự của protein kháng nguyên HIV. Mỗi nhóm chứa sáu con thỏ.

Bảng 1: Chế độ trị liệu vaccine được thử nghiệm trong nghiên cứu khả năng gây miễn dịch ở thỏ

Nhóm	Gây miễn dịch lần thứ nhất và gây miễn dịch lần thứ hai			Gây miễn dịch lần thứ ba và gây miễn dịch lần thứ tư			N=
	vectơ adeno	Liều lượng (vp)	Liều lượng tổng cộng (vp)	thúc đẩy protein	Liều lượng (ug)	Tá dược	
1	rAd26.Mos1Env	2,5x10 ¹⁰	5x10 ¹⁰	GP140 (nhánh C)	10	AdjuPhos 250µg	6
	rAd26.Mos1GagPol	1,25x10 ¹⁰					
	rAd26.Mos2Gagpol	1,25x10 ¹⁰					

2	rAd26.Mos1Env	1,25x10 ¹⁰	5x10 ¹⁰	GP140 (nhánh C)	10	AdjuPhos 250µg	6
	rAd26.C4D7	1,25x10 ¹⁰					
	rAd26.Mos1GagPol	1,25x10 ¹⁰					
	rAd26.Mos2Gagpol	1,25x10 ¹⁰					
3	rAd26.Mos1Env	1,25x10 ¹⁰	5x10 ¹⁰	GP140 (nhánh C)	10	AdjuPhos 250µg	6
	rAd26.sC4	1,25x10 ¹⁰					
	rAd26.Mos1GagPol	1,25x10 ¹⁰					
	rAd26.Mos2Gagpol	1,25x10 ¹⁰					
đối chứng	rAd26.Empty	5x10 ¹⁰	5x10 ¹⁰	Không có	0	AdjuPhos 250µg	6

So sánh vaccin Ad26 ba thành phần (thiếu kháng nguyên Env mới theo sáng chế) với vaccin Ad26 bốn thành phần (mà có chứa kháng nguyên sC4 hoặc C4D7 mới) cho phép thử nghiệm xem kháng nguyên mới có góp phần vào độ rộng của sự bảo vệ hay không. Thử nghiệm trung hòa dựa trên tế bào TZM-bl đã được thiết lập [Montefiori DC. *Methods Mol Biol* 2009,485:395-405; Sarzotti-Kelsoe M et al., *J Immunol Methods* 2014,409:131-146] được sử dụng để đo hoạt tính làm trung hòa của các ứng viên vaccin.

Kết quả được thể hiện trên Fig. 7, và được phân tích về mặt thống kê bằng cách sử dụng vaccin ba thành phần (nhóm 1 trong Bảng 1) làm nhóm đối chứng và so sánh với mỗi vaccin bốn thành phần (các nhóm 2 và 3 trong Bảng 1).

Nhìn chung, có nguồn gốc từ C4 mới (tức là mã hoá protein Env có chứa SEQ ID NO: 8, là phương án thay thế đối với mos2Env) có khả năng gây miễn dịch sau hai lần gây miễn dịch trong cơ cùng loại ở thỏ.

Khả năng trung hòa của huyết thanh miễn dịch thỏ chống lại pseudovirut Bậc 1B không có mặt (dữ liệu không được thể hiện), là điều không được mong đợi khi đã biết rằng các virut này là khó trung hòa hơn.

Khả năng trung hòa pseudovirut của huyết thanh miễn dịch thỏ chống lại virut Bậc 1A nhánh B không bị ảnh hưởng bởi sự bổ sung của các thành phần mới (dữ liệu không được thể hiện). Điều này chứng minh rằng kháng nguyên mới không gây trở ngại bất lợi tới khả năng gây miễn dịch của kháng nguyên nhánh B hiện tại có mặt trong vaccin (mặc dù các thành phần mới được định hướng đến nhánh C, sự cản trở không mong muốn này có thể không được loại trừ theo cách tiên nghiệm trước khi nó được thử nghiệm).

Khả năng trung hòa pseudovirut của huyết thanh miễn dịch thử chống lại virut Bậc 1A nhánh C được tăng cường đáng kể trong C4D7 mới có bốn thành phần chứa adeno (bốn thành phần, nhóm 2), khi so với một mình sự gây miễn dịch ba thành phần (chỉ có mos1Env) (nhóm 1) (Fig. 7 khung B). Ngoài ra, khả năng trung hòa pseudovirut của huyết thanh miễn dịch thử chống lại virut Bậc 1A nhánh C ở tuần 8 được tăng cường đáng kể trong sC4 mới có bốn thành phần chứa adeno (bốn thành phần, nhóm 3), khi so với sự gây miễn dịch ba thành phần (chỉ có mos1Env) một mình (nhóm 1) (Fig. 7 khung B).

Tóm lại, các cấu trúc C4D7 và sC4 được mã hóa trong Ad26 có khả năng gây miễn dịch và việc bổ sung chúng làm mở rộng khả năng liên kết và trung hòa của vacxin mà có mos1Env (chủ yếu nhánh B) làm thành phần Env được mã hóa Ad26 duy nhất, với các chủng nhánh C (Fig. 7B).

Ví dụ 6: Khả năng gây miễn dịch của chế độ trị liệu vacxin bao gồm vector mã hoá protein vỏ HIV tổng hợp theo sáng chế

Một nghiên cứu thử khác đánh giá sự kết hợp vector bốn thành phần Ad26.Mos4.HIV (gồm có bốn vector adenovirut: Ad26.Mos1GagPol [mã hoá SEQ ID NO: 28], Ad26.Mos2GagPol [mã hoá SEQ ID NO: 29], Ad26.Mos1Env [mã hoá SEQ ID NO: 5] và Ad26.Mos2SEnv [tên gọi “C4D7” như được dùng ở trên cũng được đề cập đến dưới dạng “Mos2S”; vector này mã hoá SEQ ID NO: 18 mới theo sáng chế], trong hỗn hợp 1:1:1:1 ở liều lượng tổng cộng 5×10^9 vp,) được dùng trong cơ dưới dạng sự gây miễn dịch mỗi kíp ở tuần 0 và 6, ở dạng kết hợp với sự thúc đẩy bằng protein Env HIV-1 tái tổ hợp bằng cách sử dụng gp140 Nhánh C [có trình tự gồm các gốc axit amin 30-708 của SEQ ID NO: 7], gp140 khảm [có trình tự gồm các gốc axit amin 30-724 của SEQ ID NO: 36], hoặc sự kết hợp của gp140 Nhánh C và gp140 khảm, ở tuần 13 và 19. Việc thúc đẩy bằng protein này được dùng trong cơ ở liều lượng tổng cộng bằng 10 hoặc 50 microgam của protein kết hợp với 250 mcg tá dược nhôm phosphat đã được tạo chế phẩm vào ngày gây miễn dịch.

Kết quả chỉ ra rằng tất cả các chế độ được thử nghiệm có khả năng gây miễn dịch ở tất cả các con vật, gây ra độ chuẩn kháng thể cao và hoạt tính làm trung hòa trung bình chống lại virut kiểu giả Env Bậc 1. Nếu gp140 Khảm được sử dụng làm kháng nguyên vacxin, một mình hoặc kết hợp với gp140 Nhánh C, độ chuẩn ELISA đặc hiệu gp140 Khảm và sự nhận diện pseudovirut Nhánh B được tăng lên đáng kể ở tuần 15 khi so sánh với nhóm tham chiếu được thúc đẩy bằng gp140 Nhánh C một mình. Quy mô tác dụng tổng cộng của sự cải thiện là trung bình, và lớn hơn đối với nhóm được thúc đẩy bằng dạng kết hợp gp140 Nhánh C — gp140 Khảm hai thành phần khi so với một mình gp140 Khảm. Ở tuần 21 của nghiên cứu, sự khác nhau này mất đi và đáp ứng miễn

dịch được đo đối với các đoàn hệ tiếp nhận sự thúc đẩy bằng gp140 Nhánh C — gp140 Khảm hai thành phần hoặc sự thúc đẩy bằng gp140 Nhánh C một thành phần là không thể phân biệt được về mặt thống kê.

Chế độ trị liệu protein hai thành phần cho thấy sự cảm ứng có thể so sánh được của độ chuẩn ELISA Nhánh C và sự nhận diện pseudovirus như chế độ trị liệu được thúc đẩy bằng một mình gp140 Nhánh C, cho thấy rằng sự bao gồm của chất gây miễn dịch liên quan đến nhánh B gp140 Khảm không có tác dụng bất lợi lên mức độ phủ kháng nguyên nhánh C, mặc dù làm tăng cường đáng kể mức độ phủ nhánh B ở tuần 15 của nghiên cứu.

Dữ liệu này xác nhận rằng vectơ Ad26.Mos2SEnv mã hoá kháng nguyên Env tổng hợp theo sáng chế có thể được sử dụng thành công trong chế độ trị liệu vaccin.

TÀI LIỆU THAM KHẢO

1. Barouch et al, *Nat Med* 2010, 16: 319-323
2. WO 2010/059732
3. Schiernle et al., *PNAS* 94: 8640-8645, 1997
4. Abrahamyan et al., *J Virol* 79: 106-115, 2005
5. US20120076812
6. Barouch et al., *Cell* 155:1-9, 2013
7. Havenga, et al., 2006, *J Gen Virol* 87: 2135-43;
8. WO 03/104467
9. WO 2004/001032
10. WO 2007/104792
11. Abbink *et al.*, (2007) *Virol* 81(9): 4654-63
12. Bằng sáng chế Mỹ số 7,270,811
13. Vogels *et al.*, (2003) *J Virol* 77(15): 8263-71
14. WO 00/70071
15. WO2012/082918
16. Walker LM, Phogat SK, Chan-Hui PY, Wagner D, Phung P, Goss JL, *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* 2009,326:285-289.
17. Haynes BF, Gilbert PB, McElrath MJ, Zolla-Pazner S, Tomaras GD, Alam SM, *et al.* Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N Engl J Med* 2012,366:1275-1286.
18. Barouch et al. (2015) *Science* 349: 320-324
19. Montefiori DC. Measuring HIV neutralization in a luciferase reporter gene assay. *Methods Mol Biol* 2009,485:395-405.

20. Sarzotti-Kelsoe M, Bailer RT, Turk E, Lin CL, Bilska M, Greene KM, *et al.* Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. *J Immunol Methods* 2014,409:131-146.
21. Edwards et al., *J. Virology*, 2002, 76:2683-2691.

Yêu cầu bảo hộ

1. Axit nucleic mã hoá protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 8, hoặc SEQ ID NO:8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (i) I529P, (ii) K480E, và (iii) sự kết hợp của EK479-480RRRR, I529P, A471C và T575C.
2. Axit nucleic theo điểm 1, trong đó protein vỏ HIV tổng hợp còn chứa trình tự tín hiệu, ví dụ, trình tự tín hiệu có chứa trình tự axit amin được chọn từ nhóm gồm có từ SEQ ID NO:9 đến SEQ ID NO: 12.
3. Axit nucleic theo điểm 1 hoặc 2, trong đó protein vỏ HIV tổng hợp còn chứa vùng xuyên màng, ví dụ, vùng xuyên màng có chứa SEQ ID NO:13, tốt hơn là protein vỏ HIV tổng hợp còn chứa SEQ ID NO:37 dung hợp với đầu tận cùng C của SEQ ID NO:8 và đầu tận cùng N của vùng xuyên màng.
4. Axit nucleic theo điểm 3, trong đó protein vỏ HIV tổng hợp còn chứa mảnh của vùng tế bào chất, tốt hơn là mảnh của vùng tế bào chất có chứa trình tự axit amin được nêu trong SEQ ID NO:14 hoặc các gốc 1-4 của chúng.
5. Axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 4 nêu trên, trong đó protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO: 18.
6. Axit nucleic theo điểm 1 hoặc 2, trong đó protein vỏ HIV tổng hợp: (a) còn chứa vùng trime hoá, ví dụ vùng trime hoá có trình tự axit amin được nêu trong SEQ ID NO: 15 hoặc SEQ ID NO:16; hoặc (b) có chứa SEQ ID NO:8 có tổ hợp các đột biến EK479-480RRRR, I529P, A471C và T575C.
7. Axit nucleic theo điểm 6, trong đó protein vỏ HIV tổng hợp có chứa trình tự axit amin gồm các gốc 1-686 của SEQ ID NO: 19.
8. Vector có chứa axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 7, trong đó axit nucleic này được liên kết có điều khiển với trình tự khởi động.
9. Vector theo điểm 8, là vector virus, tốt hơn là vector adenovirus.

10. Vector theo điểm 9, trong đó vector adenovirut là vector typ huyết thanh adenovirut người 26 (Ad26).

11. Tế bào được phân lập có chứa vector theo điểm bất kỳ trong số các điểm từ 8 đến 10.

12. Chế phẩm có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector theo điểm bất kỳ trong số các điểm từ 8 đến 10, và chất mang.

13. Vacxin kết hợp, bao gồm:

- (i) chế phẩm thứ nhất có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector adenovirut, tốt hơn là vector adenovirut 26, mã hoá protein vỏ HIV tổng hợp có trình tự axit amin được nêu trong SEQ ID NO: 8, tốt hơn là SEQ ID NO: 18;
- (ii) chế phẩm thứ hai có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của vector adenovirut thứ hai, tốt hơn là vector adenovirut 26 thứ hai, mã hoá polypeptit kháng nguyên HIV có chứa trình tự axit amin được nêu trong SEQ ID NO: 5; và tùy ý
- (iii) ít nhất một chế phẩm bổ sung có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của ít nhất một vector được chọn từ nhóm gồm có
 - (iiia) vector mã hoá ít nhất một polypeptit kháng nguyên có trình tự axit amin được chọn từ nhóm gồm có các SEQ ID NO: 1-4, 28 và 29, và
 - (iiib) polypeptit có chứa lượng hữu hiệu để gây đáp ứng miễn dịch của polypeptit kháng nguyên HIV được phân lập có các gốc 30-708 của trình tự axit amin được nêu trong SEQ ID NO: 7, hoặc các gốc 30-724 của SEQ ID NO: 36,

trong đó chế phẩm thứ nhất, chế phẩm thứ hai và chế phẩm bổ sung có mặt trong cùng chế phẩm hoặc trong một hoặc nhiều chế phẩm khác nhau.

14. Protein vỏ HIV tổng hợp có chứa trình tự axit amin được nêu trong SEQ ID NO:8 hoặc SEQ ID NO:8 có một hoặc nhiều đột biến được chọn từ nhóm gồm có (i) I529P, (ii) K480E, và (iii) sự kết hợp của EK479-480RRRR, I529P, A471C và T575C, tốt hơn là, có chứa (i) SEQ ID NO:8 có tổ hợp các đột biến EK479-480RRRR, I529P, A471C

và T575C; (ii) các gốc axit amin 30-704 của SEQ ID NO: 18, hoặc (iii) các gốc axit amin 30-686 của SEQ ID NO: 19.

15. Vector adenovirut 26 mã hóa protein vỏ HIV tổng hợp chứa hoặc gồm có trình tự axit amin được nêu trong SEQ ID NO: 18.

16. Chế phẩm chứa vector adenovirut 26 mã hóa protein vỏ HIV tổng hợp chứa trình tự axit amin được nêu trong SEQ ID NO: 18.

17. Vector MVA mã hóa protein vỏ HIV tổng hợp chứa hoặc gồm có trình tự axit amin được nêu trong SEQ ID NO: 18.

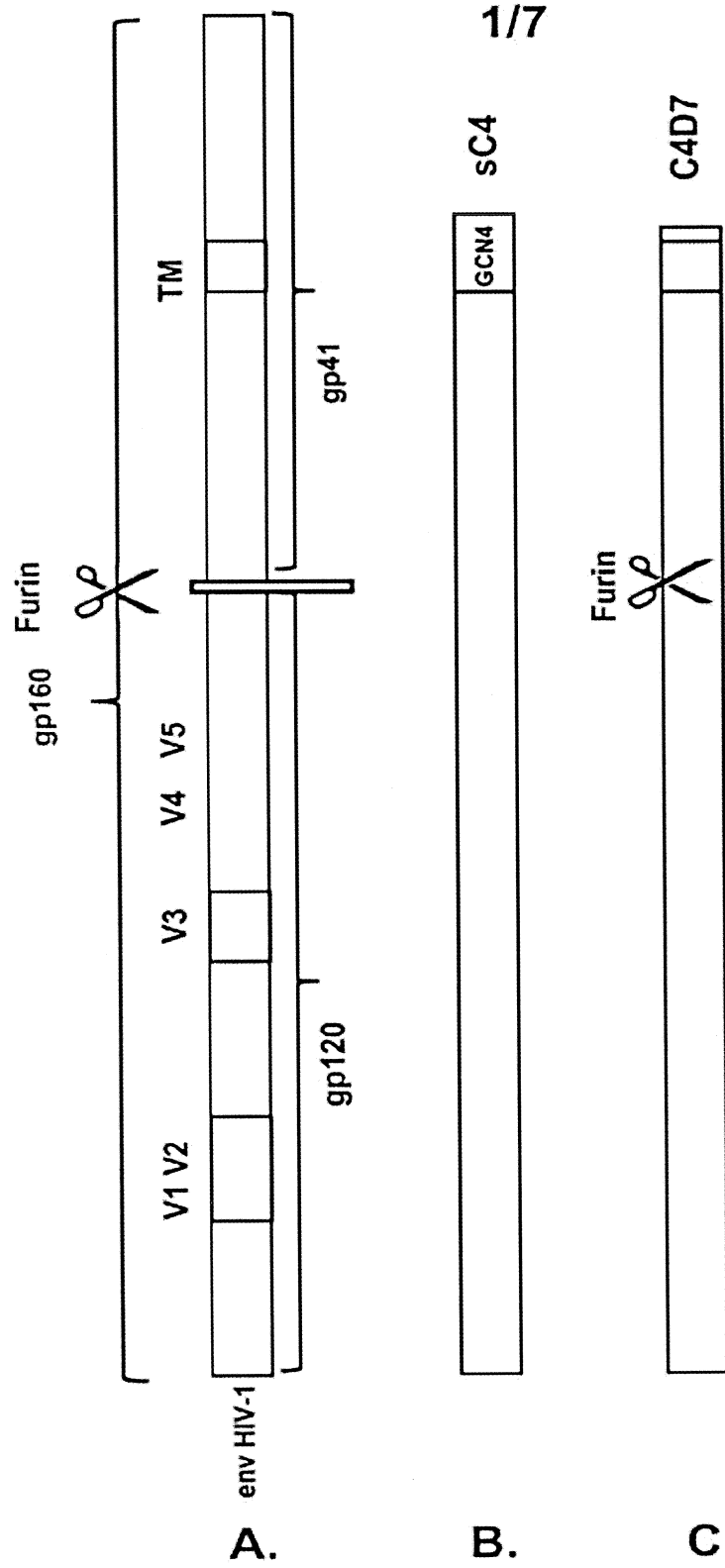


Fig. 1

2/7

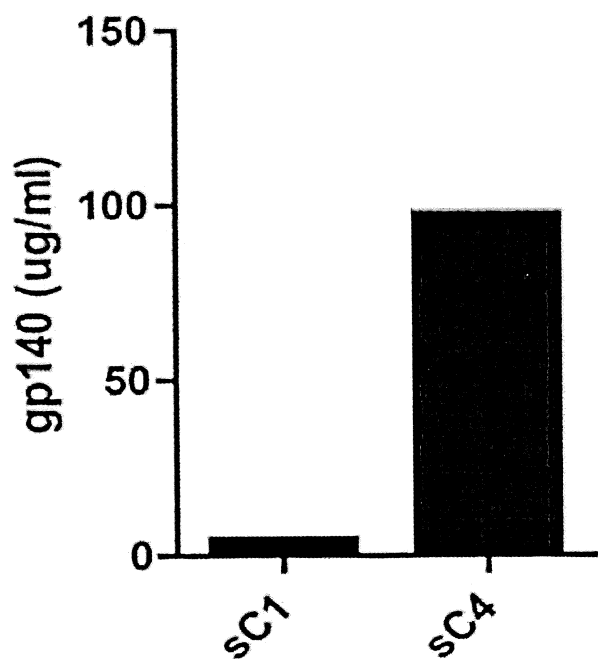


Fig. 2

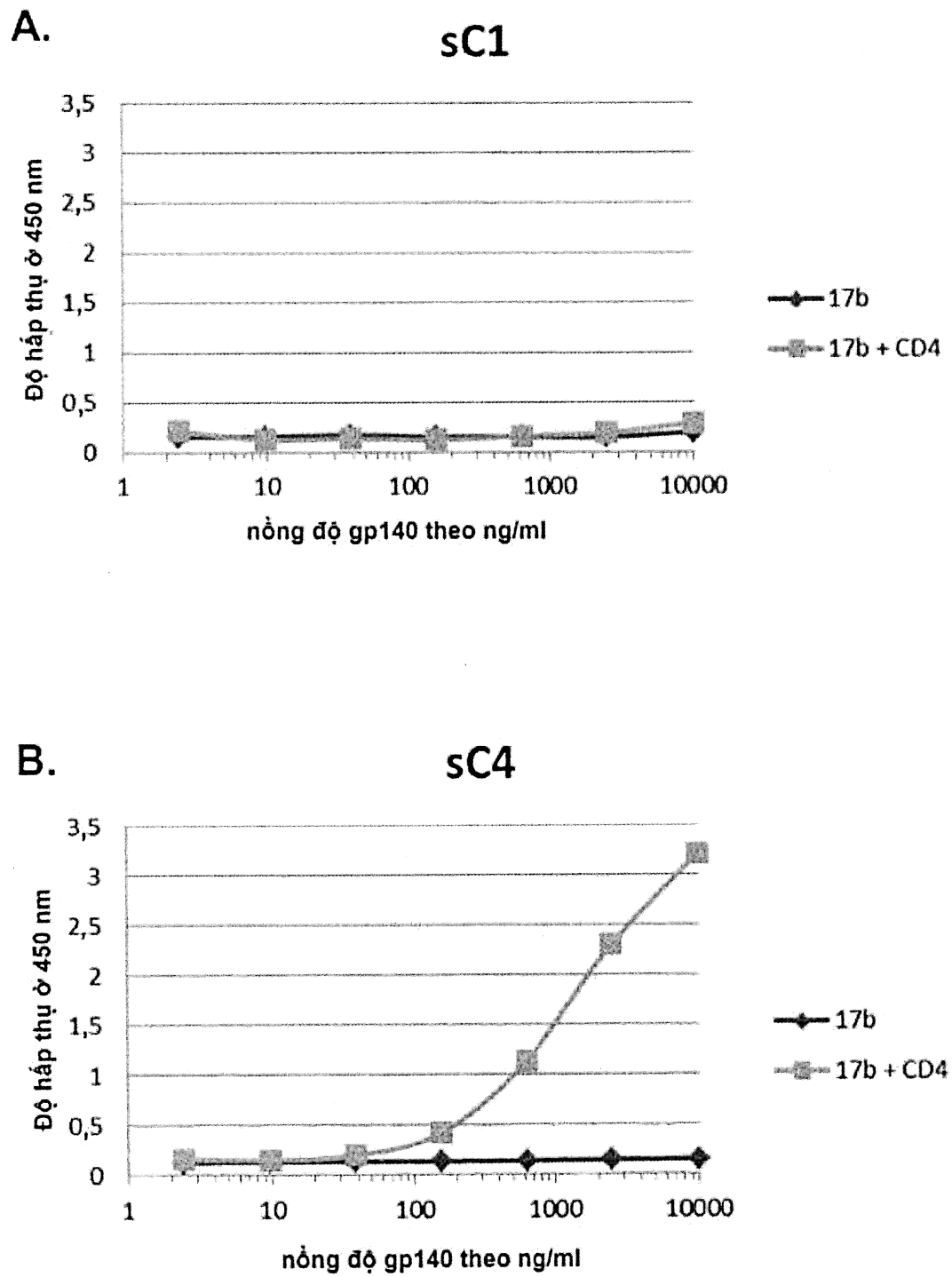


Fig. 3

4/7

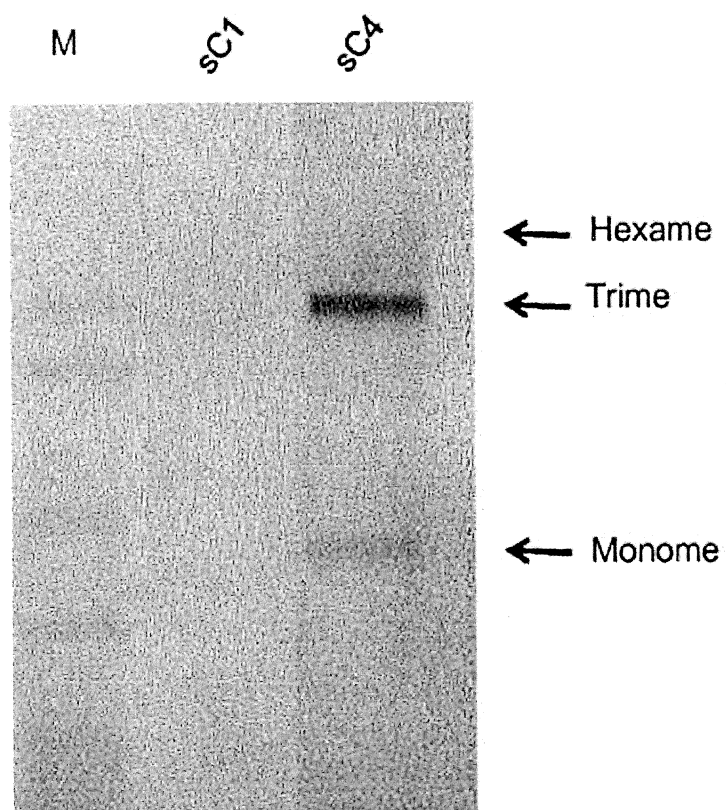


Fig. 4

5/7

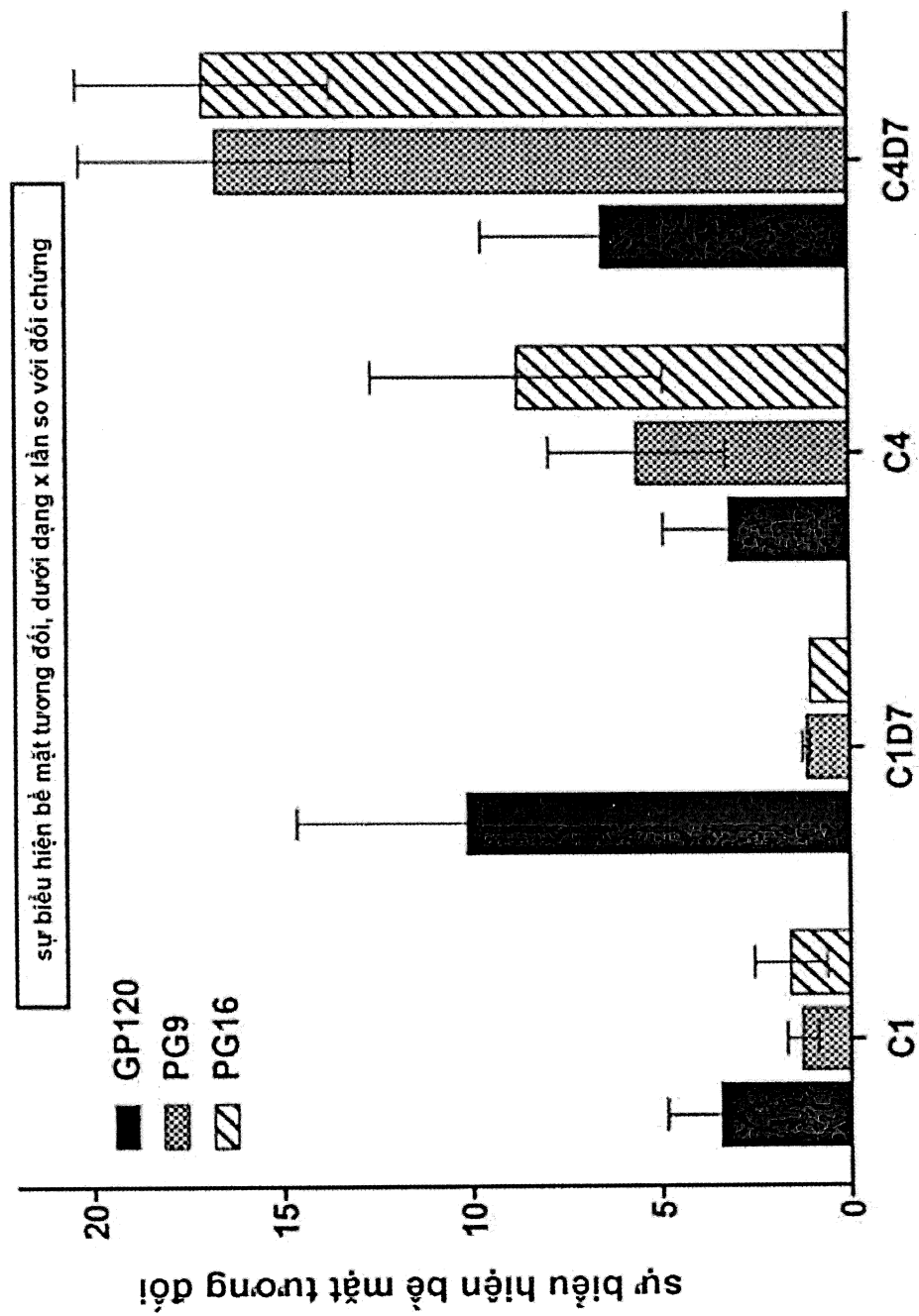


Fig. 5

6/7

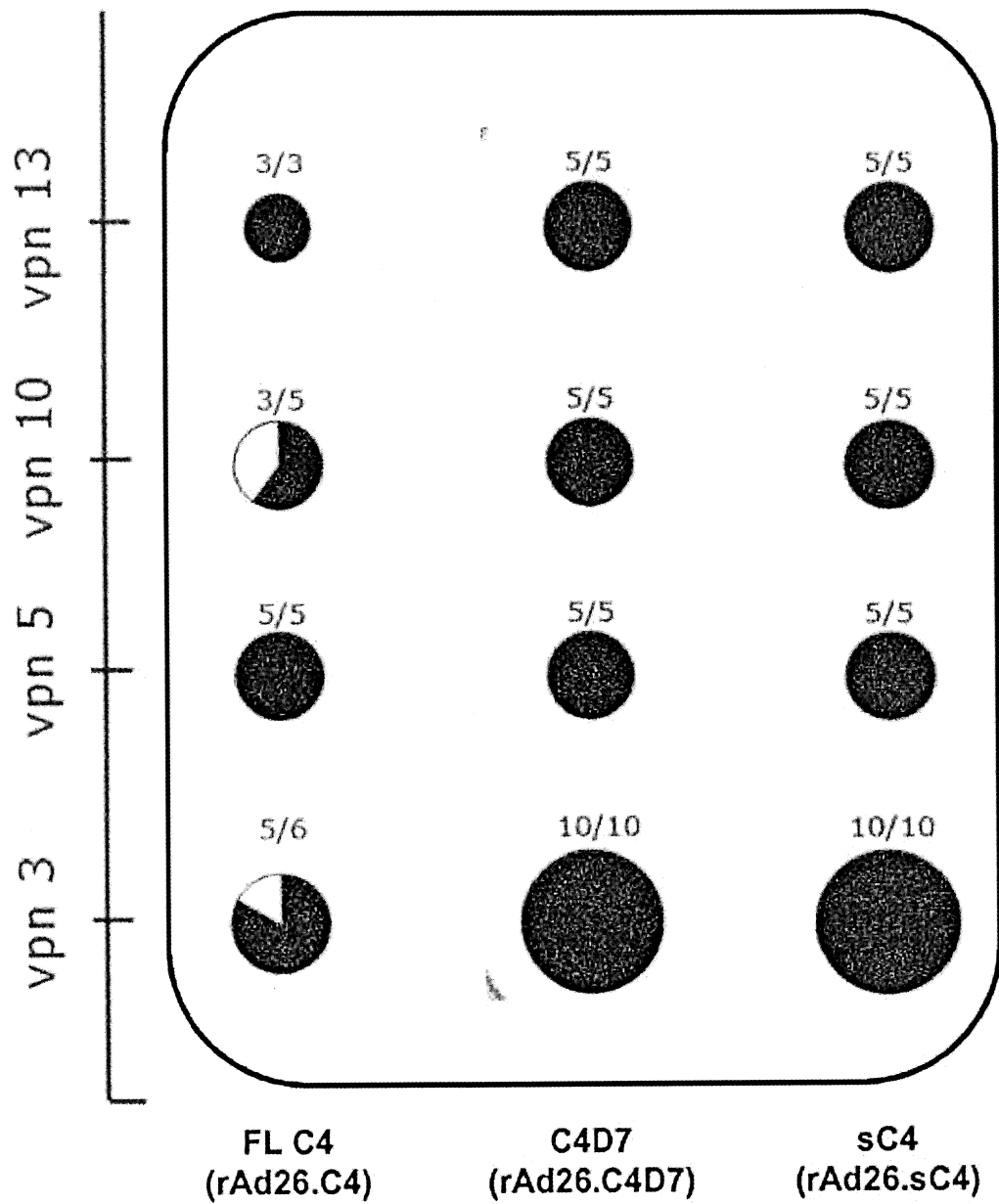


Fig. 6

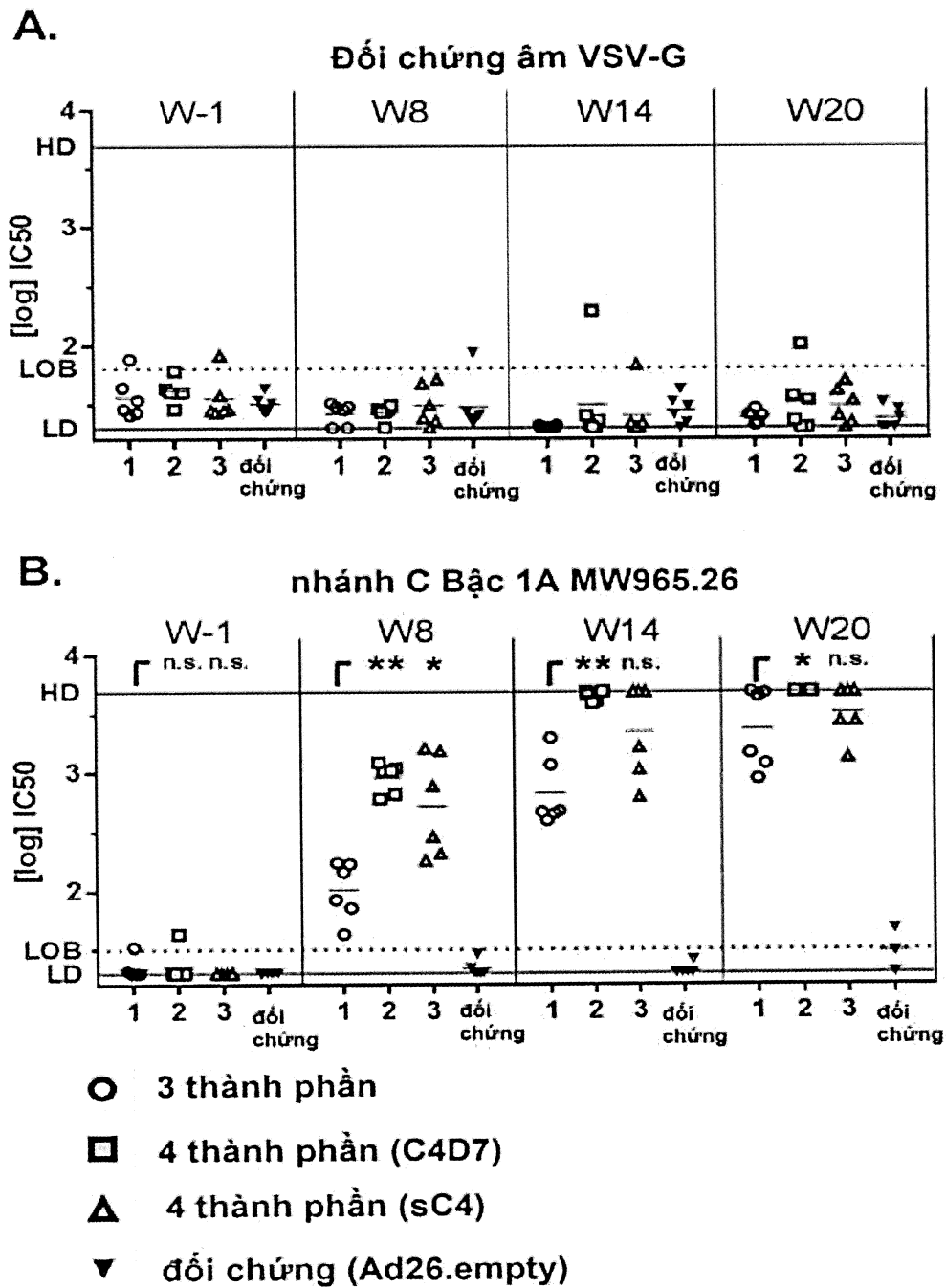


Fig. 7

DANH MỤC TRÌNH TỰ

<110> Janssen Vaccines & Prevention B.V.

<120> AXIT NUCLEIC MÃ HOÁ PROTEIN VỎ HIV TỔNG HỢP, VECTO CHỨA AXIT NUCLEIC, CHẾ PHẨM VÀ VACCIN BAO GỒM CHẾ PHẨM NÀY

<130> 0265 EP P01 PRI

<160> 37

<170> PatentIn phiên bản 3.5

<210> 1

<211> 500

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự kháng nguyên khảm mos1Gag

<400> 1

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu

50 55 60

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp

85 90 95

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

100 105 110

Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Gln Val

115 120 125

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

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

<210> 2

<211> 491

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự kháng nguyên khảm mos2Gag

<400> 2

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Lys Trp
 1 5 10 15
 Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys
 20 25 30

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

Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys Glu
 275 280 285
 Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu
 290 295 300
 Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr Leu Leu Val
 305 310 315 320
 Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro
 325 330 335
 Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
 340 345 350
 Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Thr Asn
 355 360 365
 Ser Thr Ile Leu Met Gln Arg Ser Asn Phe Lys Gly Ser Lys Arg Ile
 370 375 380
 Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn Cys
 385 390 395 400
 Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His
 405 410 415
 Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile
 420 425 430
 Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro
 435 440 445
 Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr Thr
 450 455 460
 Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro Leu Thr Ser Leu
 465 470 475 480
 Arg Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln
 485 490
 <210> 3
 <211> 850

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự kháng nguyên khảm mos1Pol

<400> 3

```

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

```

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

435	440	445	
Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asp Arg Gly			
450	455	460	
Arg Gln Lys Ile Val Ser Leu Thr Glu Thr Thr Asn Gln Lys Thr Ala			
465	470	475	480
Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Ser Glu Val Asn			
485	490	495	
Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro			
500	505	510	
Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile			
515	520	525	
Lys Lys Glu Arg Val Tyr Leu Ser Trp Val Pro Ala His Lys Gly Ile			
530	535	540	
Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ser Gly Ile Arg Lys			
545	550	555	560
Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His Glu Lys			
565	570	575	
Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro			
580	585	590	
Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp Gln Cys Gln Leu Lys			
595	600	605	
Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln			
610	615	620	
Leu Ala Cys Thr His Leu Glu Gly Lys Ile Ile Leu Val Ala Val His			
625	630	635	640
Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly			
645	650	655	
Gln Glu Thr Ala Tyr Phe Ile Leu Lys Leu Ala Gly Arg Trp Pro Val			
660	665	670	
Lys Val Ile His Thr Ala Asn Gly Ser Asn Phe Thr Ser Ala Ala Val			

675 680 685
 Lys Ala Ala Cys Trp Trp Ala Gly Ile Gln Gln Glu Phe Gly Ile Pro
 690 695 700
 Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Met Asn Lys Glu Leu
 705 710 715 720
 Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr
 725 730 735
 Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly
 740 745 750
 Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Ile Asp Ile Ile Ala Thr
 755 760 765
 Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Ile Lys Ile Gln Asn
 770 775 780
 Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro Ile Trp Lys Gly Pro
 785 790 795 800
 Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn
 805 810 815
 Ser Asp Ile Lys Val Val Pro Arg Arg Lys Val Lys Ile Ile Lys Asp
 820 825 830
 Tyr Gly Lys Gln Met Ala Gly Ala Asp Cys Val Ala Gly Arg Gln Asp
 835 840 845

Glu Asp

850

<210> 4

<211> 850

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự kháng nguyên khảm mos2Pol

<400> 4

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

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

465 470 475 480
 Leu Gln Ala Ile His Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn
 485 490 495
 Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro
 500 505 510
 Asp Lys Ser Glu Ser Glu Leu Val Ser Gln Ile Ile Glu Gln Leu Ile
 515 520 525
 Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile
 530 535 540
 Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Arg Gly Ile Arg Lys
 545 550 555 560
 Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His Glu Lys
 565 570 575
 Tyr His Ser Asn Trp Arg Ala Met Ala Ser Glu Phe Asn Leu Pro Pro
 580 585 590
 Ile Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys
 595 600 605
 Gly Glu Ala Ile His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln
 610 615 620
 Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His
 625 630 635 640
 Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly
 645 650 655
 Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val
 660 665 670
 Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Ser Ala Thr Val
 675 680 685
 Lys Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro
 690 695 700
 Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Ile Asn Lys Glu Leu

705 710 715 720
 Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr
 725 730 735
 Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly
 740 745 750
 Ile Gly Glu Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Ser
 755 760 765
 Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn
 770 775 780
 Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro Leu Trp Lys Gly Pro
 785 790 795 800
 Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn
 805 810 815
 Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp
 820 825 830
 Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp
 835 840 845
 Glu Asp
 850

<210> 5

<211> 685

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự kháng nguyên khảm mos1Env

<400> 5

Met Arg Val Thr Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp
 1 5 10 15
 Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala Ala Gly Lys

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

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

500 505 510
 Ser Glu Lys Ser Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu
 515 520 525
 Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val
 530 535 540
 Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu
 545 550 555 560
 Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp
 565 570 575
 Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu
 580 585 590
 Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile
 595 600 605
 Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu
 610 615 620
 Asp Lys Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile
 625 630 635 640
 Asn Asn Tyr Thr Ser Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn
 645 650 655
 Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala
 660 665 670
 Ser Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp Leu Trp
 675 680 685

<210> 6

<211> 684

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự kháng nguyên khảm mos2Env

<400> 6

```

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

```

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

465 470 475 480
 Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys
 485 490 495
 Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Ser
 500 505 510
 Glu Lys Ser Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly
 515 520 525
 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
 530 535 540
 Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
 545 550 555 560
 Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
 565 570 575
 Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln
 580 585 590
 Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys
 595 600 605
 Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr
 610 615 620
 Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly
 625 630 635 640
 Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
 645 650 655
 Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn
 660 665 670
 Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp
 675 680

<210> 7

<211> 708

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trime gp140 nhánh C được làm ổn định: C97ZA012-gp140-foldon với các đột biến phân cắt

<400> 7

Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp

1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Asn

20 25 30

Leu Trp Val Gly Asn Met Trp Val Thr Val Tyr Tyr Gly Val Pro Val

35 40 45

Trp Thr Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Thr Lys Ala

50 55 60

Tyr Asp Arg Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro

65 70 75 80

Thr Asp Pro Asn Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn

85 90 95

Phe Asn Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile

100 105 110

Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro

115 120 125

Leu Cys Val Thr Leu His Cys Thr Asn Ala Thr Phe Lys Asn Asn Val

130 135 140

Thr Asn Asp Met Asn Lys Glu Ile Arg Asn Cys Ser Phe Asn Thr Thr

145 150 155 160

Thr Glu Ile Arg Asp Lys Lys Gln Gln Gly Tyr Ala Leu Phe Tyr Arg

165 170 175

Pro Asp Ile Val Leu Leu Lys Glu Asn Arg Asn Asn Ser Asn Asn Ser

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

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

660 665 670
 Ser Arg Ile Glu Gly Arg Gly Ser Gly Gly Tyr Ile Pro Glu Ala Pro
 675 680 685
 Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu
 690 695 700
 Ser Thr Phe Leu
 705
 <210> 8
 <211> 625
 <212> PRT
 <213> Trình tự Nhân tạo
 <220>
 <223> mảnh C4: gp41 bị cắt cụt gp120 không có peptit tín hiệu và vùng xuyên
 màng
 <400> 8
 Met Gly Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
 1 5 10 15
 Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu
 20 25 30
 Lys Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
 35 40 45
 Pro Asn Pro Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn
 50 55 60
 Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser
 65 70 75 80
 Leu Trp Asp Ala Ser Leu Glu Pro Cys Val Lys Leu Thr Pro Leu Cys
 85 90 95
 Val Thr Leu Asn Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr
 100 105 110
 Tyr Asn Ile Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser

115	120	125
Phe Asn Ala Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala		
130	135	140
Leu Phe Tyr Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser		
145	150	155
Glu Lys Ser Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys		
165	170	175
Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro		
180	185	190
Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys		
195	200	205
Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr		
210	215	220
Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu		
225	230	235
Leu Asn Gly Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn		
245	250	255
Leu Thr Asn Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val		
260	265	270
Asn Ile Thr Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg		
275	280	285
Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp		
290	295	300
Ile Arg Gln Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr		
305	310	315
Leu Gln Gly Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr		
325	330	335
Ile Lys Phe Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His		
340	345	350
Thr Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu		

```

355              360              365
Phe Asn Glu Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro
370              375              380
Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala
385              390              395              400
Ile Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile
              405              410              415
Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro
              420              425              430
Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn
              435              440              445
Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu
              450              455              460
Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Lys
465              470              475              480
Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala
              485              490              495
Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg
              500              505              510
Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala
              515              520              525
Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys
              530              535              540
Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln
545              550              555              560
Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr
              565              570              575
Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile
              580              585              590
Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr

```

595 600 605
 Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu
 610 615 620

Lys

625

<210> 9

<211> 29

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự tín hiệu

<400> 9

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp

1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val

20 25

<210> 10

<211> 29

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự tín hiệu

<400> 10

Met Arg Val Thr Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp

1 5 10 15

Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala

20 25

<210> 11

<211> 29

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự tín hiệu

<400> 11

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp

1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val

20 25

<210> 12

<211> 29

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự tín hiệu

<400> 12

Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp

1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala

20 25

<210> 13

<211> 22

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> vùng xuyên màng

<400> 13

Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe

1 5 10 15

Ala Val Leu Ser Ile Val

20

<210> 14

<211> 7

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> vùng tế bào chất bị cắt cụt

<400> 14

Asn Arg Val Arg Gln Gly Tyr

1 5

<210> 15

<211> 32

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> miền trime hóa GCN4

<400> 15

Met Lys Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile Tyr

1 5 10 15

His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu Val

20 25 30

<210> 16

<211> 30

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> miền trime hóa foldon

<400> 16

Gly Ser Gly Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr

1 5 10 15

Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu

20 25 30

<210> 17

<211> 862

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự C4

<400> 17

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp

1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val Met Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

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

115 120 125

Asn Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile

130 135 140

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

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu Phe Asn Glu
 385 390 395 400
 Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro Cys Arg Ile
 405 410 415
 Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Ile Tyr Ala
 420 425 430
 Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu
 435 440 445
 Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro Asn Asp Thr
 450 455 460
 Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn Trp Arg Ser
 465 470 475 480
 Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu Gly Val Ala
 485 490 495
 Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Lys Arg Ala Val
 500 505 510
 Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala Gly Ser Thr
 515 520 525
 Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu
 530 535 540
 Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala
 545 550 555 560
 Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln
 565 570 575
 Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln Gln Leu Leu
 580 585 590
 Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro
 595 600 605
 Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asp Asn
 610 615 620

Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr Thr Gly Glu
 625 630 635 640
 Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu
 645 650 655
 Lys Asp Leu Leu Ala Leu Asp Ser Trp Asn Asn Leu Trp Asn Trp Phe
 660 665 670
 Ser Ile Ser Lys Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val
 675 680 685
 Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val
 690 695 700
 Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Leu Gln Thr Leu Thr
 705 710 715 720
 Gln Asn Pro Gly Gly Leu Asp Arg Leu Gly Arg Ile Glu Glu Glu Gly
 725 730 735
 Gly Glu Gln Asp Lys Asp Arg Ser Ile Arg Leu Val Asn Gly Phe Phe
 740 745 750
 Ala Leu Phe Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His
 755 760 765
 Arg Leu Arg Asp Phe Ile Leu Ile Val Ala Arg Ala Val Glu Leu Leu
 770 775 780
 Gly Arg Ser Ser Leu Arg Gly Leu Gln Arg Gly Trp Glu Ile Leu Lys
 785 790 795 800
 Tyr Leu Gly Ser Leu Leu Gln Tyr Trp Gly Leu Glu Leu Lys Lys Ser
 805 810 815
 Ala Ile Asn Leu Leu Asp Thr Ile Ala Ile Ala Val Ala Glu Gly Thr
 820 825 830
 Asp Arg Ile Ile Glu Leu Ile Gln Arg Ile Cys Arg Ala Ile Cys Asn
 835 840 845
 Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala Ala Leu Gln
 850 855 860

<210> 18

<211> 711

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự C4D7

<400> 18

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp

1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val Met Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

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

115 120 125

Asn Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile

130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala

145 150 155 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr

165 170 175

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

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

Lys Asp Leu Leu Ala Leu Asp Ser Trp Asn Asn Leu Trp Asn Trp Phe

660

665

670

Ser Ile Ser Lys Trp Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val

675

680

685

Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala Val Leu Ser Ile Val

690

695

700

Asn Arg Val Arg Gln Gly Tyr

705

710

<210> 19

<211> 704

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự sC4

<400> 19

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp

1

5

10

15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val Met Gly Asn

20

25

30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Lys

35

40

45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val

50

55

60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65

70

75

80

Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85

90

95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100

105

110

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

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Lys Phe
 355 360 365
 Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His Thr Phe Asn
 370 375 380
 Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu Phe Asn Glu
 385 390 395 400
 Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro Cys Arg Ile
 405 410 415
 Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Ile Tyr Ala
 420 425 430
 Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu
 435 440 445
 Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro Asn Asp Thr
 450 455 460
 Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn Trp Arg Ser
 465 470 475 480
 Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu Gly Val Ala
 485 490 495
 Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Glu Arg Ala Val
 500 505 510
 Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala Gly Ser Thr
 515 520 525
 Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu
 530 535 540
 Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala
 545 550 555 560
 Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln
 565 570 575
 Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln Gln Leu Leu
 580 585 590

Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro

595

600

605

Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asp Asn

610

615

620

Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr Thr Gly Glu

625

630

635

640

Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Met Lys

645

650

655

Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile Tyr His Ile

660

665

670

Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu Val Gly Ser

675

680

685

Gly Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg

690

695

700

<210> 20

<211> 4050

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho mos1GagPol

<400> 20

atgggagcca gagccagcgt gctgtccgga ggggagctgg accgctggga gaagatcagg	60
ctgaggcctg gaggaagaa gaagtacagg ctgaagcaca tcgtgtgggc cagcagagag	120
ctggaacggt ttgccgtgaa ccctggcctg ctggaaacca gcgagggtg taggcagatt	180
ctgggacagc tgcagcccag cctgcagaca ggcagcgagg aactgcggag cctgtacaac	240
accgtggcca ccctgtactg cgtgcaccag cggatcgaga tcaaggacac caaagaagcc	300
ctggaaaaga tcgaggaaga gcagaacaag agcaagaaga aagcccagca ggctgccgct	360
gacacaggca acagcagcca ggtgtcccag aactacccca tcgtgcagaa catccaggga	420
cagatgggtgc accaggccat cagccctcgg accctgaacg cctgggtgaa ggtggtggag	480

gaaaaggcct tcagccctga ggtgatcccc atgttctctg ccctgagcga gggagccaca	540
ccccaggacc tgaacaccat gctgaacacc gtgggagggc accaggctgc catgcagatg	600
ctgaaagaga caatcaacga ggaagctgcc gagtgggaca gggtcacccc agtgcacgct	660
ggacctatcg ctcttgcca gatgagagag cccagaggca gcgatattgc tggcaccacc	720
tccacactgc aggaacagat cggctggatg accaacaacc ctcccatccc tgtgggagag	780
atctacaagc ggtggatcat tctgggactg aacaagatcg tgcggatgta cagccctgtg	840
agcatcctgg acatcaggca gggacccaaa gagcccttca gggactacgt ggaccggttc	900
tacaagaccc tgagagccga gcaggccagc caggacgtga agaactggat gaccgagaca	960
ctgctggtgc agaacgcaa ccctgactgc aagaccatcc tgaaagccct gggacctgct	1020
gccaccctgg aagagatgat gacagcctgc cagggagtgaggaggacctgg ccacaaggcc	1080
agggtgctgg ccgaggccat gagccaggtg accaactctg ccaccatcat gatgcagaga	1140
ggcaacttcc ggaaccagag aaagaccgtg aagtgcttca actgtggcaa agagggacac	1200
attgccaaga actgcagggc tcccaggaag aaaggctgct ggaagtgcgg aaaagaaggc	1260
caccagatga aggactgcac cgagaggcag gccaaacttcc tgggcaagat ctggcctagc	1320
aacaagggca ggcctggcaa ctctctgcag aacagaccgg agcccaccgc tcctcccgag	1380
gaaagcttcc ggtttggcga ggaaccacc acccctagcc agaagcagga acccatcgac	1440
aaagagatgt accctctggc cagcctgaag agcctgttcg gcaacgaccc cagcagccag	1500
atggctccca tcagcccaat cgagacagtg cctgtgaagc tgaagcctgg catggacgga	1560
cccagggtga agcagtggcc tctgaccgag gaaaagatca aagccctgac agccatctgc	1620
gaggaaatgg aaaaagaggg caagatcacc aagatcggac ccgagaaccc ctacaacacc	1680
cctgtgttcg ccatcaagaa gaaagacagc accaagtgga ggaaactggt ggacttcaga	1740
gagctgaaca agcggaccca ggacttctgg gaggtgcagc tgggcatccc tcaccctgct	1800
ggcctgaaga aaaagaaaag cgtgaccgtg ctggctgtgg gagatgccta cttcagcgtg	1860
cctctggacg agggcttcg gaagtacaca gccttcacca tcccagcac caacaacgag	1920
acacctggca tcagatacca gtacaacgtg ctgcctcagg gctggaaagg cagccctgcc	1980
atcttccagt gcagcatgac cagaatcctg gaacccttca gagccaagaa ccctgagatc	2040
gtgatctacc agtatatggc tgccctctac gtgggcagcg acctggaaat cggacagcac	2100
agagccaaaa tcgaagaact ccgcgagcac ctgctgaagt ggggattcac caccctgac	2160
aagaagcacc agaaagagcc tcccttctg tggatgggct acgagctgca ccctgacaag	2220
tggaccgtgc agcccatcca gctgccagag aaggactcct ggaccgtgaa cgacatccag	2280

aaactggtcg gcaagctgaa ctgggccagc cagatctacc ctggcatcaa agtcagacag	2340
ctgtgtaagc tgctgagggg agccaaagca ctgaccgaca tcgtgcctct gacagaagaa	2400
gccgagctgg aactggccga gaacagagag atcctgaaag aaccctgca cggagtgtac	2460
tacgaccct ccaaggacct gattgccgag atccagaaac agggacacga ccagtggacc	2520
taccagatct atcaggaacc tttcaagaac ctgaaaacag gcaagtacgc caagatgcgg	2580
acagcccaca ccaacgacgt gaagcagctg accgaagccg tgcagaaaat cgccatggaa	2640
agcatcgtga tctggggaaa gacaccaag ttcaggctgc ccatccagaa agagacatgg	2700
gaaacctggt ggaccgacta ctggcaggcc acctggattc ccgagtggga gttcgtgaac	2760
acccacccc tgggtgaagct gtggtatcag ctggaaaagg accctatcgc tggcgtggag	2820
acattctacg tggctggagc tgccaacaga gagacaaagc tgggcaaggc tggctacgtg	2880
accgacagag gcagacagaa aatcgtgagc ctgaccgaaa ccaccaacca gaaaacagcc	2940
ctgcaggcca tctatctggc actgcaggac agcgggaagc aggtgaacat cgtgacagcc	3000
agccagtatg ccctgggcat catccaggcc cagcctgaca agagcgagag cgagctggtg	3060
aaccagatca tcgagcagct gatcaagaaa gaacgggtgt acctgagctg ggtgccagcc	3120
cacaagggca tcggagggaa cgagcaggtg gacaagctgg tgtccagcgg aatccggaag	3180
gtgctgttcc tggacggcat cgataaagcc caggaagagc acgagaagta ccacagcaat	3240
tggagagcca tggccagcga cttcaacctg cctcccgtgg tggccaaaga aatcgtggcc	3300
agctgcgacc agtgccagct gaaaggcgag gccatgcacg gacaggtgga ctgctcccct	3360
ggcatctggc agctggcatg caccacactg gaaggcaaga tcattctggt ggccgtgcac	3420
gtggccagcg gatacatcga agccgaagtg atccctgccg agacaggga ggaacagcc	3480
tacttcatcc tgaagctggc tggcagatgg cctgtgaagg tgatccacac agccaacggc	3540
agcaacttca cctctgctgc cgtgaaggct gcctgttggg gggctggcat tcagcaggaa	3600
tttggcatcc cctacaatcc ccagtctcag ggagtgggtg ccagcatgaa caaagagctg	3660
aagaagatca tcggacaggt cagggatcag gccgagcacc tgaaaactgc cgtccagatg	3720
gccgtgttca tccacaactt caagcggaag ggagggatcg gaggtactc tgctggcgag	3780
cggatcatcg acatcattgc caccgatatc cagaccaaag agctgcagaa acagatcatc	3840
aagatccaga acttcagggt gtactacagg gacagcaggg acccatctg gaagggacct	3900
gccaagctgc tgtggaaagg cgaaggagcc gtcgtcatcc aggacaacag cgacatcaag	3960
gtggtgccca gacggaaggt gaaaatcatc aaggactacg gcaaacagat ggctggagcc	4020
gactgtgtcg ctggcaggca ggacaggac	4050

<210> 21

<211> 4023

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho mos2GagPol

<400> 21

atgggagcca gagccagcat cctgcgagga ggggaagctgg acaagtggga gaagatcagg	60
ctgaggcctg gagggaagaa acactacatg ctgaagcacc tggctctgggc cagcagagag	120
ctggaacggt ttgccctcaa tcctggcctg ctggaaacca gcgagggctg caagcagatc	180
atcaagcagc tgcagcctgc cctgcagaca ggcaccgagg aactgcggag cctgttcaac	240
accgtggcca ccctgtactg cgtgcatgcc gagatcgaag tgagggacac caaagaagcc	300
ctggacaaga tcgaggaaga gcagaacaag agccagcaga aaaccagca ggccaaagaa	360
gccgacggca aggtctccca gaactacccc atcgtgcaga acctgcaggg acagatggtg	420
caccagccca tcagccctcg gacactgaat gcctgggtga aggtgatcga ggaaaaggcc	480
ttcagccctg aggtgatccc catgttcaca gccctgagcg agggagccac accccaggac	540
ctgaacacca tgctgaacac cgtgggaggg caccaggctg ccatgcagat gctgaaggac	600
accatcaacg aggaagctgc cgagtgggac aggctgcacc ctgtgcacgc tggacctgtg	660
gctcctggcc agatgagaga gccagaggc agcgatattg ctggcaccac ctccaatctg	720
caggaacaga tcgcctggat gaccagcaac cctcccatcc ctgtgggaga catctacaag	780
cggtggatca tcctgggact gaacaagatc gtgcggatgt acagccctac ctccatcctg	840
gacatcaagc agggacccaa agagcctttc agggactacg tggaccggtt cttcaagacc	900
ctgagagccg agcaggccac ccaggacgtg aagaactgga tgaccgacac cctgctggtg	960
cagaacgcca accctgactg caagaccatc ctgagagccc tgggacctgg agccaccctg	1020
gaagagatga tgacagcctg ccaggggagt ggaggaccct ctcacaaggc tagggtgctg	1080
gccgaggcca tgagccagac caacagcacc atcctgatgc agcgagcaa cttcaagggc	1140
agcaagcgga tcgtgaagtg cttcaactgt ggcaaagagg gacacattgc cagaaactgt	1200
agggcaccca ggaagaaagg ctgctggaag tgcggaaaag aaggccacca gatgaaggac	1260
tgcaccgaga ggcaggccaa cttcctgggc aagatctggc ctagccacaa gggcagacct	1320

ggcaacttcc tgcagagcag acccgagccc accgctcctc cagccgagag cttccggttc	1380
gaggaaacca ccctgctcc caagcaggaa cctaaggaca gagagcctct gaccagcctg	1440
agaagcctgt tcggcagcga ccctctgagc cagatggctc ccatctcccc tatcgagaca	1500
gtgcctgtga agctgaagcc tggcatggac ggacccaagg tgaaacagtg gcctctgacc	1560
gaggaaaaga tcaaagccct ggtggagatc tgtaccgaga tggaaaaaga gggcaagatc	1620
agcaagatcg gacccgagaa cccctacaac acccctatct tcgccatcaa gaagaaagac	1680
agcaccaagt ggaggaaact ggtggacttc agagagctga acaagcggac ccaggacttc	1740
tgggaggtgc agctgggcat ccctcaccct gctggcctga agaaaaagaa aagcgtgacc	1800
gtgctggccg tgggagatgc ctacttcagc gtgcctctgg acgaggactt cagaaagtac	1860
acagccttca ccatccccag catcaacaac gagacacctg gcatcagata ccagtacaac	1920
gtgctgcctc agggatggaa gggctctcct gcaatcttcc agagcagcat gaccaagatc	1980
ctggaaccct tccggaagca gaaccctgac atcgtgatct accagtacat ggcagccctg	2040
tacgtcggca gcgacctgga aatcggacag caccggacca agatcgaaga actcaggcag	2100
cacctgctgc ggtggggatt caccaccctt gacaagaagc accagaaaga gcctcccttc	2160
ctgtggatgg gctacgagct gcacccagac aagtggaccg tgcagcccat cgtgctgcct	2220
gagaaggact cctggaccgt gaacgacatc cagaaactgg tcggcaagct gaactgggcc	2280
agccagatct acgctggcat caaagtgaag cagctgtgta agctcctgag aggcacaaa	2340
gccctgaccg aggtggtgcc actgacagag gaagccgagc tggaactggc cgagaacaga	2400
gagatcctga aagaaccctg gcacggagtg tactacgacc ccagcaagga cctgattgcc	2460
gagatccaga agcagggaca gggacagtgg acctaccaga tctaccagga acccttcaag	2520
aacctgaaaa caggcaagta cgccaggatg aggggagccc acaccaacga cgtcaaacag	2580
ctgaccgaag ccgtgcagaa gatcgccacc gagagcatcg tgatttgggg aaagacaccc	2640
aagttcaagc tgcccatcca gaaagagaca tgggaggcct ggtggaccga gtactggcag	2700
gccacctgga ttcccgagtg ggagttcgtg aacacccac ccctggtgaa gctgtggtat	2760
cagctggaaa aagaacccat cgtgggagcc gagacattct acgtggctgg agctgccaac	2820
agagagacaa agctgggcaa ggctggctac gtgaccgaca gaggcaggca gaaagtggtg	2880
tcctgaccg ataccaccaa ccagaaaaca gccctgcagg ccatccacct ggctctgcag	2940
gactctggcc tggaagtga catcgtgaca gccagccagt atgccctggg catcattcag	3000
gcacagcctg acaagagcga gagcgagctg gtgtctcaga tcattgagca gctgatcaag	3060
aaagaaaagg tgtacctggc ctgggtgcca gccacaagg ggatcggagg gaacgagcag	3120

gtggacaagc tgggtgtccag gggcatccgg aaggtgctgt ttctggacgg catcgacaaa 3180
 gcccaggaag agcacgagaa gtaccacagc aattggagag ccatggccag cgagttcaac 3240
 ctgcctccca tcgtggccaa agaaatcgtg gcctcttgcg acaagtgcc a gctgaaaggc 3300
 gaggccattc acggacaggt ggactgcagc ccaggcatct ggcagctggc ctgcacccac 3360
 ctggaaggca aggtgatcct ggtggccgtg cacgtggcct ctggatacat cgaagccgaa 3420
 gtgatccctg ccgagacagg ccaggaaaca gcctacttcc tgctgaagct ggctggcagg 3480
 tggcctgtga aaaccatcca cacagccaac ggcagcaact tcacctctgc caccgtgaag 3540
 gctgcctgtt ggtgggctgg cattaagcag gaatttgga tcccctacaa ccctcagtct 3600
 caggggagtgg tggcctccat caacaaagag ctgaagaaga tcatcgga ggtcagggat 3660
 caggccgagc atctgaaaac agccgtccag atggccgtgt tcatccacaa cttcaagcgg 3720
 aaggaggagg tccgagagta ctctgctggc gagaggatcg tggacattat cgccagcgat 3780
 atccagacca aagaactgca gaagcagatc acaaagatcc agaacttcag ggtgtactac 3840
 agggacagca gagatcccct gtggaaggga cctgccaaagc tgctgtggaa aggcgaagga 3900
 gccgtcgtca tccaggacaa cagcgacatc aaggtggtgc ccagacggaa ggccaagatc 3960
 atcagagact acggcaaaca gatggctggc gacgactgcg tcgcctctag gcaggacgag 4020
 gac 4023

<210> 22

<211> 2055

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho mos1Env

<400> 22

atgcgggtga ccggcatccg gaagaactac cagcacctgt ggcggtgggg caccatgctg 60
 ctgggcatcc tgatgatttg ctctgccgcc ggaaagctgt gggtgaccgt gtactacggc 120
 gtgcccgtgt ggaaagaggc caccaccacc ctgttctgcg ccagcgacgc caaggcctac 180
 gacaccgagg tgcacaacgt gtgggccacc cacgcctgcg tgcccaccga cccaacccc 240
 caggaagtgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa caacatggtg 300
 gagcagatgc acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcgtgaag 360

ctgaccccc	tgtgctgac	cctgaactgc	accgacgacg	tgcggaacgt	gaccaacaac	420
gccaccaaca	ccaacagcag	ctggggcgag	cctatggaaa	agggcgagat	caagaactgc	480
agcttcaaca	tcaccacctc	catccggaac	aaggtgcaga	agcagtacgc	cctgttctac	540
aagctggacg	tggtgccc	cgacaacgac	agcaacaaca	ccaactaccg	gctgatcagc	600
tgcaacacca	gcgtgatcac	ccaggcctgc	ccaaggtgt	ccttcgagcc	catccccatc	660
cactactgcg	cccctgccgg	cttcgccatc	ctgaagtgca	acgacaagaa	gttcaacggc	720
accggcccc	gcaccaacgt	gagcaccgtg	cagtgcaccc	acggcatccg	gcccgtggtg	780
tccaccacgc	tgctgctgaa	cggcagcctg	gccgaggaag	aggtggtgat	cagaagcgag	840
aatttcacca	acaatgccaa	gaccatcatg	gtgcagctga	acgtgagcgt	ggagatcaac	900
tgcaccggc	ccaacaacaa	cacccggaag	agcatccaca	tcggccctgg	cagggccttc	960
tacacagccg	gcgacatcat	cggcgacatc	cggcaggccc	actgcaacat	cagccgggcc	1020
aactggaaca	acaccctgcg	gcagatcgtg	gagaagctgg	gcaagcagtt	cggcaacaac	1080
aagaccatcg	tgttcaacca	cagcagcggc	ggagaccccg	agatcgtgat	gcacagcttc	1140
aactgtggcg	gcgagttctt	ctactgcaac	agcaccaagc	tgttcaacag	cacctggacc	1200
tggaacaact	ccacctggaa	taacaccaag	cggagcaacg	acaccgaaga	gcacatcacc	1260
ctgccctgcc	ggatcaagca	gattatcaat	atgtggcagg	aggtcggcaa	ggccatgtac	1320
gcccctccca	tccggggcca	gatccggtgc	agcagcaaca	tcaccggcct	gctgctgacc	1380
cgggacggcg	gcaacgatac	cagcggcacc	gagatcttcc	ggcctggcgg	cggagatatg	1440
cgggacaact	ggcggagcga	gctgtacaag	tacaaggtgg	tgaagatcga	gcccctgggc	1500
gtggctccca	ccaaggccaa	gcggcgggtg	gtgcagagcg	agaagagcgc	cgtgggcatc	1560
ggcgccgtgt	ttctgggctt	cctgggagcc	gccggaagca	ccatgggagc	cgccagcatg	1620
accctgaccg	tgcaggcccc	gctgctgctg	tccggcatcg	tgcagcagca	gaacaacctg	1680
ctccgggcca	tcgaggccca	gcagcacctg	ctgcagctga	ccgtgtgggg	catcaagcag	1740
ctgcaggcca	gggtgctggc	cgtggagaga	tacctgaagg	atcagcagct	cctggggatc	1800
tggggctgca	gcggcaagct	gatctgcacc	accaccgtgc	cctggaacgc	cagctgggtcc	1860
aacaagagcc	tggacaagat	ctggaacaat	atgacctgga	tggaatggga	gcgcgagatc	1920
aacaattaca	ccagcctgat	ctacaccctg	atcgaggaaa	gccagaacca	gcaggaaaag	1980
aacgagcagg	aactgctgga	actggacaag	tgggccagcc	tgtggaactg	gttcgacatc	2040
agcaactggc	tgtgg					2055

<210> 23

<211> 2052

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho mos2Env

<400> 23

atgagagtgc ggggcatcca gcggaactgg cccagtggt ggatctgggg catcctgggc	60
ttttggatga tcatcatctg ccgggtgatg ggcaacctgt gggtgaccgt gtactacggc	120
gtgcccgtgt ggaaagaggc caagaccacc ctgttctgcg ccagcgacgc caaggcctac	180
gagaaagagg tgcacaacgt gtggggccacc cacgcctgcg tgcccaccga ccccaacccc	240
caggaaatgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg	300
gaccagatgc acgaggacat catccggctg tgggaccaga gcctgaagcc ctgctggaag	360
ctgaccccc tgtgctgac cctggaatgc cggaactga gaaactgag cagcaacggc	420
acctacaaca tcatccaca cgagacctac aaagagatga agaactgcag cttcaacgcc	480
accaccgtgg tggaggaccg gaagcagaag gtgcacgccc tgttctaccg gctggacatc	540
gtgcccctgg acgagaacaa cagcagcgag aagtccagcg agaacagctc cgagtactac	600
cggctgatca actgcaacac cagcgccatc acccaggcct gcccgaaggt gtccttcgac	660
cccatcccca tccactactg cgcccctgcc ggctacgcca tcctgaagtg caacaacaag	720
accttcaacg gcaccggccc ctgcaacaac gtgagcaccg tgcagtgcac ccacggcatc	780
aagcccgtgg tgtccacca gctgctgctg aacggcagcc tggccgagga agagatcatc	840
atccggtccg agaacctgac caacaacgcc aagaccatca tcgtgcacct gaatgagacc	900
gtgaacatca cctgcacccg gcccacaac aacaccggga agagcatccg gatcggccct	960
ggccagacct tttagccac cggcgacatc atcgggcgaca tccggcaggc cactgcaac	1020
ctgagccggg acggctggaa caagaccctg cagggcggtga agaagaagct ggccgagcac	1080
ttccccaata agaccatcaa cttcaccagc agcagcgcg gagacctgga aatcaccacc	1140
cacagcttca actgcagggg cgagttcttc tactgcaata cctccggcct gttcaatggc	1200
acctacatgc ccaacggcac caacagcaac agcagcagca acatcaccct gccctgccgg	1260
atcaagcaga tcatcaatat gtggcaggag gtcggcaggg ccatgtacgc ccctcccatc	1320
gccggcaata tcacctgccg gtccaacatc accggcctgc tgctgaccag ggacggcggc	1380

agcaacaacg gcgtgcctaa cgacaccgag accttccggc ctggcggcgg agatatgctg 1440
 aacaactggc ggagcgagct gtacaagtac aaggtggtg aggtgaagcc cctgggcgtg 1500
 gctcctaccg aggccaagcg gcgggtggtg gagagcgaga agagcgccgt gggcatcggc 1560
 gccgtgtttc tgggcattct gggagccgcc ggaagcacca tgggagccgc cagcatcacc 1620
 ctgaccgtgc aggcccggca gctgctgtcc ggcacgtgc agcagcagag caacctgctg 1680
 agagccatcg aggcccagca gcacatgctg cagctgaccg tgtggggcat caagcagctg 1740
 cagaccggg tgctggccat cgagagatac ctgcaggatc agcagctcct gggcctgtgg 1800
 ggctgcagcg gcaagctgat ctgcaccacc gccgtgccct ggaacaccag ctggtccaac 1860
 aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agagatcggc 1920
 aactacaccg gcgagatcta caggctgctg gaagagagcc agaaccagca ggaaaagaac 1980
 gagaaggacc tgctggccct ggacagctgg aagaacctgt ggaactgggt cgacatcacc 2040
 aactggctgt gg 2052

<210> 24

<211> 829

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> vùng khởi động CMV dùng để biểu hiện kháng nguyên trong vật truyền Ad26

<400> 24

tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60
 ttggccattg catacgttgt atccatatca taatatgtac atttatattg gctcatgtcc 120
 aacattaccg ccatgttgac attgattatt gactagttat taatagtaat caattacggg 180
 gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc 240
 gcctggctga ccgccaacg acccccggcc attgacgtca ataatgacgt atgttcccat 300
 agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc 360
 ccacttggca gtacatcaag tgtatcatat gccaaagtacg cccctattg acgtcaatga 420
 cggtaaattg cccgcctggc attatgccca gtacatgacc ttatgggact ttcctacttg 480
 gcagtacatc tacgtattag tcatcgctat taccatgggt atgcggtttt ggcagtacat 540

caatgggCGT ggatagCGgt ttgactcag gggatttcca agtctccacc ccattgacgt 600
 caatgggagT ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc 660
 cgccccattg acgcaaattg gcggtagggc tgtacgggtg gaggtctata taagcagagc 720
 tcgttttagtg aaccgtcaga tcgcctggag acgccatcca cgctgttttg acctccatag 780
 aagacaccgg gaccgatcca gcctccgCGg ccgggaacgg tgcattgga 829

<210> 25

<211> 2586

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho C4

<400> 25

atgagagtgc ggggcatgct gagaaactgg cagcagtggT ggatctggtc cagcctgggc 60
 ttctggatgc tgatgatcta cagcgtgatg ggcaacctgt gggtcaccgt gtactacggc 120
 gtgcccgtgt ggaaggacgc caagaccacc ctgttttgcg cctccgatgc caaggcctac 180
 gagaaagagg tgcacaacgt ctgggccacc cagcctgtgT tgcccaccga ccccaatccc 240
 caggaaatcg tcctgggcaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggTc 300
 gatcagatgc acgaggacat catctccctg tgggacgcct ccctggaacc ctgcgtgaag 360
 ctgaccctc tgtgcgtgac cctgaactgc cggaacgtgc gcaacgtgtc cagcaacggc 420
 acctacaaca tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct 480
 accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttctaccg gctggacatc 540
 gtgcccctgg acgagaacaa cagcagcgag aagtccctccg agaacagctc cgagtactac 600
 agactgatca actgcaacac cagcgccatc acccaggcct gcccgaaggT gtccttcgac 660
 cctatcccca tccactactg cgcccctgcc ggctacgcca tcctgaagtg caacaacaag 720
 accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc 780
 aagcccgtgg tgtctacca gctgctgctg aacggcagcc tggccgagga agagatcatt 840
 atcagaagcg agaacctgac caacaacgcc aaaaccatca tcgtccacct gaacgaaacc 900
 gtgaacatca cctgtaccCG gcctaacaac aacaccCGga agtccatccg gatcggcctt 960
 ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggg ccactgcaat 1020

ctgagccggg acggctggaa caagacactg cagggcgtca agaagaagct ggccgaacac	1080
ttccctaaca agactatcaa gttcgccct cactctggcg gcgacctgga aatcaccacc	1140
cacaccttca actgtcgggg cgagttcttc tactgcaata cctccaacct gttcaacgag	1200
agcaacatcg agcggaacga cagcatcatc acactgcctt gccggatcaa gcagattatc	1260
aatatgtggc aggaagtggg cagagccatc tacgcccctc caatcgccgg caacatcaca	1320
tgccggtcca atatcaccgg cctgctgctc accagagatg gcggtccaa caatggcgtg	1380
ccaaacgaca ccgagacatt cagacccggc ggaggcgaca tgcggaacaa ttggcggagc	1440
gagctgtaca agtacaaggt ggtggaagtg aagcccctgg gcggtggccc taccgaggcc	1500
aagagaagag tggtcgaacg cgagaagcgg gccgtgggaa tcggagccgt gtttctggga	1560
atcctgggag ccgctggctc taccatgggc gctgcctcta tcaccctgac agtgcaggcc	1620
agacagctgc tcagcggcat cgtgcagcag cagagcaacc tgctgagagc cattgaggcc	1680
cagcagcaca tgctgcagct gaccgtgtgg ggcattaagc agctccagac acgggtgctg	1740
gccatcgaga gatacctgca ggatcagcag ctctggggc tgtggggctg tagcggcaag	1800
ctgatctgta ccaccgccgt gccctggaat acctcttggg gcaacaagag ccagaccgac	1860
atctgggaca acatgacctg gatgcagtgg gacaaagaaa tcggcaacta taccggcgag	1920
atctatagac tgctggaaga gtcccagaac cagcaggaaa agaacgagaa ggacctgctg	1980
gccctggatt cttggaacaa tctgtggaac tggttcagca tctccaagtg gctgtggtac	2040
atcaagatct tcatcatgat cgtgggcggc ctgatcggcc tcgggatcat ctttgccgtg	2100
ctgagcatcg tgaaccgcgt gcggcaggga tacagccctc tgagcctgca gaccctgact	2160
cagaaccctg gcggactgga cagactgggc cggattgagg aagaaggcgg cgagcaggac	2220
aaggatcgga gcatcaggct ggtcaacggc ttcttcgctc tgttttgga cgacctgcgg	2280
agcctgtgcc tgttcagcta ccacagactg cgggacttta tcctgattgt ggccagagcc	2340
gtcgaactgc tggggagaag ctctctgaga ggctgcagc ggggctggga gattctgaag	2400
tacctgggct ccctgctgca gtactggggc ctggaactga agaagtctgc catcaatctg	2460
ctcgacacaa tcgctattgc cgtggccgaa ggcaccgata gaatcatcga gctgatccag	2520
cggatctgcc gggccatctg caacatcccc agacggatca gacagggctt cgaggccgct	2580
ctgcag	2586

<210> 26

<211> 2133

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho C4D7

<400> 26

atgagagtgc ggggcatgct gagaaactgg cagcagtggg ggatctgggc cagcctgggc	60
ttctggatgc tgatgatcta cagcgtgatg ggcaacctgt gggtcaccgt gtactacggc	120
gtgcccggtg ggaaggacgc caagaccacc ctgttttgcg cctccgatgc caaggcctac	180
gagaaagagg tgcacaacgt ctggggccacc cagcctgtg tgcccaccga cccaatccc	240
caggaaatcg tcctgggcaa cgtgaccgag aacttcaaca tgtggaagaa cgacatgggc	300
gatcagatgc acgaggacat catctccctg tgggacgcct ccctggaacc ctgctgaag	360
ctgacccctc tgtgctgac cctgaactgc cggaacgtgc gcaacgtgc cagcaacggc	420
acctacaaca tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct	480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttctaccg gctggacatc	540
gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac	600
agactgatca actgcaacac cagcgccatc acccaggcct gcccgaagg gtccttcgac	660
cctatcccca tccactactg cgcccctgcc ggctacgcca tcctgaagtg caacaacaag	720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc	780
aagcccgtgg tgtctacca gctgctgctg aacggcagcc tggccgagga agagatcatt	840
atcagaagcg agaacctgac caacaacgcc aaaaccatca tcgtccacct gaacgaaacc	900
gtgaacatca cctgtaccg gcctaacaac aacaccggga agtccatccg gatcggccct	960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc ccaactgcaat	1020
ctgagccggg acggctggaa caagacactg cagggcggtca agaagaagct ggccgaacac	1080
ttccctaaca agactatcaa gttagccct cactctggcg gcgacctgga aatcaccacc	1140
cacaccttca actgtcgggg cgagttcttc tactgcaata cctccaacct gttcaacgag	1200
agcaacatcg agcgggaacga cagcatcatc acactgcctt gccggatcaa gcagattatc	1260
aatatgtggc aggaagtggg cagagccatc tacgcccctc caatcgccgg caacatcaca	1320
tgccggtcca atatcaccgg cctgctgctc accagagatg gcggctccaa caatggcgtg	1380
ccaaacgaca ccgagacatt cagacccggc ggaggcgaca tgcggaacaa ttggcggagc	1440

gagctgtaca agtacaaggt ggtggaagtg aagcccctgg gcgtggcccc taccgaggcc 1500
 aagagaagag tggtcgaacg cgagaagcgg gccgtgggaa tcggagccgt gtttctggga 1560
 atcctgggag ccgctggctc taccatgggc gctgcctcta tcacctgac agtgcaggcc 1620
 agacagctgc tcagcggcat cgtgcagcag cagagcaacc tgctgagagc cattgaggcc 1680
 cagcagcaca tgctgcagct gaccgtgtgg ggcattaagc agctccagac acgggtgctg 1740
 gccatcgaga gatacctgca ggatcagcag ctccctgggcc tgtggggctg tagcggcaag 1800
 ctgatctgta ccaccgccgt gccctggaat acctcttgga gcaacaagag ccagaccgac 1860
 atctgggaca acatgacctg gatgcagtgg gacaaagaaa tcggcaacta taccggcgag 1920
 atctatagac tgctggaaga gtcccagaac cagcaggaaa agaacgagaa ggacctgctg 1980
 gccctggatt cttggaacaa tctgtggaac tggttcagca tctccaagtg gctgtggtac 2040
 atcaagatct tcatcatgat cgtggggcggc ctgatcggcc tgcggatcat ctttgccgtg 2100
 ctgagcatcg tgaaccgcgt gcggcagggc tac 2133

<210> 27

<211> 2112

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho sC4

<400> 27

atgagagtgc ggggcatgct gagaaactgg cagcagtggg ggatctggtc cagcctgggc 60
 ttctggatgc tgatgatcta cagcgtgatg ggcaacctgt gggtcaccgt gtactacggc 120
 gtgcccggtg ggaaggacgc caagaccacc ctgttttgcg cctccgatgc caaggcctac 180
 gagaaagagg tgcacaacgt ctggggccacc cagcctgtg tgcccaccga cccaatccc 240
 caggaaatcg tcctgggcaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtc 300
 gatcagatgc acgaggacat catctccctg tgggacgcct ccctggaacc ctgcgtgaag 360
 ctgaccctc tgtgcgtgac cctgaactgc cggaacgtgc gcaacgtgtc cagcaacggc 420
 acctacaaca tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct 480
 accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttctaccg gctggacatc 540
 gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac 600

```

agactgatca actgcaacac cagcgccatc acccaggcct gcccgaaggt gtccttcgac      660
cctatcccca tccactactg cgcccctgcc ggctacgcca tcctgaagtg caacaacaag      720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc      780
aagcccgtgg tgtctacca gctgctgctg aacggcagcc tggccgagga agagatcatt      840
atcagaagcg agaacctgac caacaacgcc aaaaccatca tcgtccacct gaacgaaacc      900
gtgaacatca cctgtaccog gcctaacaac aacacccgga agtccatccg gatcggccct      960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc ccactgcaat     1020
ctgagccggg acggctggaa caagacactg cagggcgctc agaagaagct ggccgaacac     1080
ttccctaaca agactatcaa gttcgcccct cactctggcg gcgacctgga aatcaccacc     1140
cacaccttca actgtcgggg cgagttcttc tactgcaata cctccaacct gttcaacgag     1200
agcaacatcg agcggaaacg cagcatcatc aactgcctt gccggatcaa gcagattatc     1260
aatatgtggc aggaagtggg cagagccatc tacgcccctc caatcgccgg caacatcaca     1320
tgccggtcca atatcaccgg cctgctgctc accagagatg gcggctccaa caatggcgtg     1380
ccaaacgaca ccgagacatt cagacccggc ggaggcgaca tgcggaacaa ttggcggagc     1440
gagctgtaca agtacaaggt ggtggaagtg aagcccctgg gcgtggcccc taccgaggcc     1500
aagagaagag tggtcgaacg cgaggaacgg gccgtgggaa tcggagccgt gtttctggga     1560
atcctgggag ccgctggctc taccatgggc gctgcctcta tcacctgac agtgcaggcc     1620
agacagctgc tcagcggcat cgtgcagcag cagagcaacc tgctgagagc cattgaggcc     1680
cagcagcaca tgctgcagct gaccgtgtgg ggcattaagc agctccagac acgggtgctg     1740
gccatcgaga gatacctgca ggatcagcag ctctggggcc tgtggggctg tagcggcaag     1800
ctgatctgta ccaccgccgt gccctggaat acctcttgga gcaacaagag ccagaccgac     1860
atctgggaca acatgacctg gatgcagtgg gacaaagaaa tcggcaacta taccggcgag     1920
atctatagac tgctggaaga gtcccagaac cagcaggaaa agatgaagca gatcgaggac     1980
aagatcgaa gattctgag caagatctac cacatcgaga acgagatcgc ccgcatcaag     2040
aaactgatcg gcgaagtggg atccggcgct cccacaaagg ccaaaagacg ggtggtgcag     2100
cgcgagaaac gc                                             2112

```

<210> 28

<211> 1350

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự kháng nguyên khảm mos1GagPol

<400> 28

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys

20 25 30

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu

50 55 60

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp

85 90 95

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

100 105 110

Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Gln Val

115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His

130 135 140

Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu

145 150 155 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser

165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly

180 185 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu

195 200 205

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

Leu Gln Asn Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg
 450 455 460
 Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp
 465 470 475 480
 Lys Glu Met Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp
 485 490 495
 Pro Ser Ser Gln Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val
 500 505 510
 Lys Leu Lys Pro Gly Met Asp Gly Pro Arg Val Lys Gln Trp Pro Leu
 515 520 525
 Thr Glu Glu Lys Ile Lys Ala Leu Thr Ala Ile Cys Glu Glu Met Glu
 530 535 540
 Lys Glu Gly Lys Ile Thr Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr
 545 550 555 560
 Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu
 565 570 575
 Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val
 580 585 590
 Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val
 595 600 605
 Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu
 610 615 620
 Gly Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Thr Asn Asn Glu
 625 630 635 640
 Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys
 645 650 655
 Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr Arg Ile Leu Glu Pro
 660 665 670
 Phe Arg Ala Lys Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Ala Ala
 675 680 685

Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys Ile
 690 695 700
 Glu Glu Leu Arg Glu His Leu Leu Lys Trp Gly Phe Thr Thr Pro Asp
 705 710 715 720
 Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu
 725 730 735
 His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys Asp
 740 745 750
 Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp
 755 760 765
 Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu
 770 775 780
 Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu Glu
 785 790 795 800
 Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val
 805 810 815
 His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln
 820 825 830
 Lys Gln Gly His Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe
 835 840 845
 Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg Thr Ala His Thr
 850 855 860
 Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Met Glu
 865 870 875 880
 Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Ile Gln
 885 890 895
 Lys Glu Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp Gln Ala Thr Trp
 900 905 910
 Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp
 915 920 925

```

Tyr Gln Leu Glu Lys Asp Pro Ile Ala Gly Val Glu Thr Phe Tyr Val
    930                      935                      940
Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val
945                      950                      955                      960
Thr Asp Arg Gly Arg Gln Lys Ile Val Ser Leu Thr Glu Thr Thr Asn
                      965                      970                      975
Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly
                      980                      985                      990
Ser Glu Val Asn Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile
                      995                      1000                      1005
Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile
    1010                      1015                      1020
Ile Glu Gln Leu Ile Lys Lys Glu Arg Val Tyr Leu Ser Trp Val
    1025                      1030                      1035
Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu
    1040                      1045                      1050
Val Ser Ser Gly Ile Arg Lys Val Leu Phe Leu Asp Gly Ile Asp
    1055                      1060                      1065
Lys Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala
    1070                      1075                      1080
Met Ala Ser Asp Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile
    1085                      1090                      1095
Val Ala Ser Cys Asp Gln Cys Gln Leu Lys Gly Glu Ala Met His
    1100                      1105                      1110
Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln Leu Ala Cys Thr
    1115                      1120                      1125
His Leu Glu Gly Lys Ile Ile Leu Val Ala Val His Val Ala Ser
    1130                      1135                      1140
Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly Gln Glu
    1145                      1150                      1155

```

Thr Ala	Tyr Phe Ile Leu Lys	Leu Ala Gly Arg Trp	Pro Val Lys
1160	1165	1170	
Val Ile	His Thr Ala Asn Gly	Ser Asn Phe Thr Ser	Ala Ala Val
1175	1180	1185	
Lys Ala	Ala Cys Trp Trp Ala	Gly Ile Gln Gln Glu	Phe Gly Ile
1190	1195	1200	
Pro Tyr	Asn Pro Gln Ser Gln	Gly Val Val Ala Ser	Met Asn Lys
1205	1210	1215	
Glu Leu	Lys Lys Ile Ile Gly	Gln Val Arg Asp Gln	Ala Glu His
1220	1225	1230	
Leu Lys	Thr Ala Val Gln Met	Ala Val Phe Ile His	Asn Phe Lys
1235	1240	1245	
Arg Lys	Gly Gly Ile Gly Gly	Tyr Ser Ala Gly Glu	Arg Ile Ile
1250	1255	1260	
Asp Ile	Ile Ala Thr Asp Ile	Gln Thr Lys Glu Leu	Gln Lys Gln
1265	1270	1275	
Ile Ile	Lys Ile Gln Asn Phe	Arg Val Tyr Tyr Arg	Asp Ser Arg
1280	1285	1290	
Asp Pro	Ile Trp Lys Gly Pro	Ala Lys Leu Leu Trp	Lys Gly Glu
1295	1300	1305	
Gly Ala	Val Val Ile Gln Asp	Asn Ser Asp Ile Lys	Val Val Pro
1310	1315	1320	
Arg Arg	Lys Val Lys Ile Ile	Lys Asp Tyr Gly Lys	Gln Met Ala
1325	1330	1335	
Gly Ala	Asp Cys Val Ala Gly	Arg Gln Asp Glu Asp	
1340	1345	1350	

<210> 29

<211> 1341

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự kháng nguyên khảm mos2GagPol

<400> 29

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Lys Trp

1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys

20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro

35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Lys Gln Leu

50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Phe Asn

65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Ala Glu Ile Glu Val Arg Asp

85 90 95

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

100 105 110

Gln Lys Thr Gln Gln Ala Lys Glu Ala Asp Gly Lys Val Ser Gln Asn

115 120 125

Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Pro Ile

130 135 140

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys Ala

145 150 155 160

Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly Ala

165 170 175

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln

180 185 190

Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu

195 200 205

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

Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr Thr
 450 455 460
 Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro Leu Thr Ser Leu
 465 470 475 480
 Arg Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln Met Ala Pro Ile Ser
 485 490 495
 Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro
 500 505 510
 Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val
 515 520 525
 Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly
 530 535 540
 Pro Glu Asn Pro Tyr Asn Thr Pro Ile Phe Ala Ile Lys Lys Lys Asp
 545 550 555 560
 Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg
 565 570 575
 Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly
 580 585 590
 Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr
 595 600 605
 Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr
 610 615 620
 Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn
 625 630 635 640
 Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser
 645 650 655
 Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val
 660 665 670
 Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile
 675 680 685

Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg
 690 695 700
 Trp Gly Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe
 705 710 715 720
 Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro
 725 730 735
 Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys
 740 745 750
 Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys
 755 760 765
 Val Lys Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu
 770 775 780
 Val Val Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg
 785 790 795 800
 Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys
 805 810 815
 Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr
 820 825 830
 Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala
 835 840 845
 Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala
 850 855 860
 Val Gln Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro
 865 870 875 880
 Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr
 885 890 895
 Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr
 900 905 910
 Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val
 915 920 925

Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys
 930 935 940
 Leu Gly Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val
 945 950 955 960
 Ser Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile His
 965 970 975
 Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser
 980 985 990
 Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser
 995 1000 1005
 Glu Leu Val Ser Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys
 1010 1015 1020
 Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn
 1025 1030 1035
 Glu Gln Val Asp Lys Leu Val Ser Arg Gly Ile Arg Lys Val Leu
 1040 1045 1050
 Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His Glu Lys Tyr
 1055 1060 1065
 His Ser Asn Trp Arg Ala Met Ala Ser Glu Phe Asn Leu Pro Pro
 1070 1075 1080
 Ile Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu
 1085 1090 1095
 Lys Gly Glu Ala Ile His Gly Gln Val Asp Cys Ser Pro Gly Ile
 1100 1105 1110
 Trp Gln Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val
 1115 1120 1125
 Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro
 1130 1135 1140
 Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala
 1145 1150 1155

Gly Arg	Trp Pro Val Lys Thr	Ile His Thr Ala Asn	Gly Ser Asn
1160	1165	1170	
Phe Thr	Ser Ala Thr Val Lys	Ala Ala Cys Trp Trp	Ala Gly Ile
1175	1180	1185	
Lys Gln	Glu Phe Gly Ile Pro	Tyr Asn Pro Gln Ser	Gln Gly Val
1190	1195	1200	
Val Ala	Ser Ile Asn Lys Glu	Leu Lys Lys Ile Ile	Gly Gln Val
1205	1210	1215	
Arg Asp	Gln Ala Glu His Leu	Lys Thr Ala Val Gln	Met Ala Val
1220	1225	1230	
Phe Ile	His Asn Phe Lys Arg	Lys Gly Gly Ile Gly	Glu Tyr Ser
1235	1240	1245	
Ala Gly	Glu Arg Ile Val Asp	Ile Ile Ala Ser Asp	Ile Gln Thr
1250	1255	1260	
Lys Glu	Leu Gln Lys Gln Ile	Thr Lys Ile Gln Asn	Phe Arg Val
1265	1270	1275	
Tyr Tyr	Arg Asp Ser Arg Asp	Pro Leu Trp Lys Gly	Pro Ala Lys
1280	1285	1290	
Leu Leu	Trp Lys Gly Glu Gly	Ala Val Val Ile Gln	Asp Asn Ser
1295	1300	1305	
Asp Ile	Lys Val Val Pro Arg	Arg Lys Ala Lys Ile	Ile Arg Asp
1310	1315	1320	
Tyr Gly	Lys Gln Met Ala Gly	Asp Asp Cys Val Ala	Ser Arg Gln
1325	1330	1335	
Asp Glu	Asp		
1340			

<210> 30

<211> 709

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự sC1

<400> 30

```

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

```

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

Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly
 450 455 460
 Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
 465 470 475 480
 Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys
 485 490 495
 Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg
 500 505 510
 Glu Glu Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly
 515 520 525
 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
 530 535 540
 Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
 545 550 555 560
 Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
 565 570 575
 Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln
 580 585 590
 Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys
 595 600 605
 Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr
 610 615 620
 Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly
 625 630 635 640
 Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
 645 650 655
 Gln Glu Lys Met Lys Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser
 660 665 670
 Lys Ile Tyr His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile
 675 680 685

Gly Glu Val Gly Ser Gly Ala Pro Thr Lys Ala Lys Arg Arg Val Val

690

695

700

Gln Arg Glu Lys Arg

705

<210> 31

<211> 867

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự C1

<400> 31

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp

1

5

10

15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val Met Gly Asn

20

25

30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys

35

40

45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val

50

55

60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65

70

75

80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85

90

95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Arg Leu Trp Asp

100

105

110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115

120

125

Glu Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile

130

135

140

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

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Gly
 385 390 395 400
 Thr Tyr Met Pro Asn Gly Thr Asn Ser Asn Ser Ser Ser Asn Ile Thr
 405 410 415
 Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly
 420 425 430
 Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser
 435 440 445
 Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly
 450 455 460
 Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
 465 470 475 480
 Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys
 485 490 495
 Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg
 500 505 510
 Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly
 515 520 525
 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
 530 535 540
 Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
 545 550 555 560
 Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
 565 570 575
 Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln
 580 585 590
 Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys
 595 600 605
 Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr
 610 615 620

Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly
 625 630 635 640
 Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
 645 650 655
 Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn
 660 665 670
 Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile
 675 680 685
 Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala
 690 695 700
 Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser
 705 710 715 720
 Leu Gln Thr Leu Thr Gln Asn Pro Gly Gly Leu Asp Arg Leu Gly Arg
 725 730 735
 Ile Glu Glu Glu Gly Gly Glu Gln Asp Lys Asp Arg Ser Ile Arg Leu
 740 745 750
 Val Asn Gly Phe Phe Ala Leu Phe Trp Asp Asp Leu Arg Ser Leu Cys
 755 760 765
 Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile Val Ala Arg
 770 775 780
 Ala Val Glu Leu Leu Gly Arg Ser Ser Leu Arg Gly Leu Gln Arg Gly
 785 790 795 800
 Trp Glu Ile Leu Lys Tyr Leu Gly Ser Leu Leu Gln Tyr Trp Gly Leu
 805 810 815
 Glu Leu Lys Lys Ser Ala Ile Asn Leu Leu Asp Thr Ile Ala Ile Ala
 820 825 830
 Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Leu Ile Gln Arg Ile Cys
 835 840 845
 Arg Ala Ile Cys Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala
 850 855 860

Ala Leu Gln

865

<210> 32

<211> 716

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự C1D7

<400> 32

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp

1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val Met Gly Asn

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Arg Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Glu Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile

130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala

145 150 155 160

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

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

Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
 645 650 655
 Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn
 660 665 670
 Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile
 675 680 685
 Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala
 690 695 700
 Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr
 705 710 715

<210> 33

<211> 2127

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho sC1

<400> 33

atgagagtgc ggggcattca gagaaactgg cccagtggt ggatctgggg catcctgggc	60
ttttggatga tcattatctg ccgcgtgatg ggcaacctgt gggtcaccgt gtactacggc	120
gtgcccgtgt ggaaagaggc caagaccacc ctgttctgcg ccagcgacgc caaggcctac	180
gagaaagagg tgcacaacgt ctggggccacc cacgcctgtg tgcccaccga ccccaatccc	240
caggaaatgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg	300
gaccagatgc acgaggacat catccggctg tgggaccaga gcctgaagcc ctgcgtgaag	360
ctgaccctc tgtgcgtgac cctggaatgc cggaacgtgc gcaacgtgtc cagcaacggc	420
acctacaata tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct	480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttttaccg gctggacatc	540
gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac	600
agactgatca actgcaacac cagcgccatc acccaggcct gcccgaaggt gtccttcgac	660
cctatcccca tccactactg cgcccctgcc ggctacgcca tcctgaagtg caacaacaag	720

accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc	780
aagcccgtgg tgtctacca gctgctgctg aacggcagcc tggccgagga agagatcatc	840
atcagaagcg agaacctgac caacaacgcc aagacaatca tcgtccacct gaacgaaacc	900
gtgaacatca cctgtacccg gcctaacaac aacacccgga agtccatccg gatcggccct	960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc ccactgcaat	1020
ctgagccggg acggctggaa caagacactg cagggcgctca agaagaagct ggccgaacac	1080
ttccccaaca aaaccatcaa cttcaccagc tcctctggcg gcgacctgga aatcaccacc	1140
cacagcttta actgcagagg cgagttcttc tactgcaata cctccggcct gttcaatgga	1200
acctacatgc ccaacgggac caacagcaac tccagcagca atatcaccct gccttgccgg	1260
atcaagcaga ttatcaatat gtggcaggaa gtgggcagag ctatgtacgc ccctccaatc	1320
gccggcaaca tcacatgcag aagcaacatt accggcctgc tgctcaccag ggacggcggc	1380
tctaacaatg gcgtgcaaaa cgacaccgag acattcagac ccggcggagg cgacatgcgg	1440
aacaattggc ggagcgagct gtacaagtac aaggtggtgg aagtgaagcc cctgggcgtg	1500
gcccctaccg aagccaagag aagagtggtc gaacgcgagg aacgggccgt gggcattgga	1560
gccgtgtttc tgggaatcct gggagccgct ggcagcacca tgggcgctgc ctctatcaca	1620
ctgacagtgc aggccagaca gctcctgagc ggcacgtgc agcagcagag caacctgctg	1680
agagccatcg aggcacagca gcacatgctg cagctgaccg tgtggggcat taagcagctc	1740
cagacacggg tgctggccat tgagagatac ctgcaggatc agcagctgct cggcctgtgg	1800
ggctgtagcg gcaagctgat ctgtaccacc gccgtgcctt ggaacacctc ctggtccaac	1860
aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agaaatcggc	1920
aactataccg gcgagatcta ccgactgctg gaagagtccc agaaccagca ggaaaagatg	1980
aagcagatcg aggacaagat cgaagagatt ctgagcaaaa tctaccacat cgagaacgag	2040
atcgcccgca tcaagaaact gatcggcgaa gtgggatccg gcgctccac aaaggccaaa	2100
agacgggtgg tgcagcgaga gaaacgc	2127

<210> 34

<211> 2601

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho C1

<400> 34

atgagagtgc ggggcattca gagaaactgg ccccagtggt ggatctgggg catcctgggc	60
ttttggatga tcattatctg ccgcgtgatg ggcaacctgt gggtcaccgt gtactacggc	120
gtgcccgtgt ggaaagaggc caagaccacc ctgttctgcy ccagcgacgc caaggcctac	180
gagaaagagg tgcacaacgt ctggggccacc cacgcctgtg tgcccaccga cccaatccc	240
caggaaatgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg	300
gaccagatgc acgaggacat catccggctg tgggaccaga gcctgaagcc ctgcytgaa	360
ctgaccctc tgtgcytgac cctggaatgc cggaacgtgc gcaacgtgcy cagcaacggc	420
acctacaata tcattccaaa cgagacatac aaagagatga agaactgcag cttcaacgct	480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttttaccg gctggacatc	540
gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac	600
agactgatca actgcaacac cagcgccatc acccaggcct gcccgaaggt gtccttcgac	660
cctatcccca tccactactg cggccctgcc ggctacgcca tcctgaagtg caacaacaag	720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc	780
aagcccgtgg tgtctacca gctgctgctg aacggcagcc tggccgagga agagatcatc	840
atcagaagcg agaacctgac caacaacgcc aagacaatca tcgtccacct gaacgaaacc	900
gtgaacatca cctgtacccg gcctaacaac aacaccgga agtccatccg gatcggccct	960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc cactgcaat	1020
ctgagccggg acggctgga caagacactg cagggcgtca agaagaagct ggccgaacac	1080
ttccccaaca aaaccatcaa cttcaccagc tcctctggcg gcgacctgga aatcaccacc	1140
cacagcttta actgcagagg cgagttcttc tactgcaata cctccggcct gttcaatgga	1200
acctacatgc ccaacgggac caacagcaac tccagcagca atatcaccct gccttgccg	1260
atcaagcaga ttatcaatat gtggcaggaa gtgggcagag ctatgtacgc ccctccaatc	1320
gccggcaaca tcacatgcag aagcaacatt accggcctgc tgctcaccag ggacggcggc	1380
tctaacaatg gcgtgcaaaa cgacaccgag acattcagac ccggcgagg cgacatgcg	1440
aacaattggc ggagcgagct gtacaagtac aaggtggtgg aagtgaagcc cctgggcgtg	1500
gcccctaccg aagccaagag aagagtgtgc gaacgcgaga agcgggccgt gggcattgga	1560
gccgtgtttc tgggaatcct gggagccgct ggcagcacca tgggcgctgc ctctatcaca	1620
ctgacagtgc agccagaca gtcctgagc ggcacgtgc agcagcagag caacctgctg	1680

agagccatcg aggcacagca gcacatgctg cagctgaccg tgtggggcat taagcagctc 1740
 cagacacggg tgctggccat tgagagatac ctgcaggatc agcagctgct cggcctgtgg 1800
 ggctgtagcg gcaagctgat ctgtaccacc gccgtgcctt ggaacacctc ctggtccaac 1860
 aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agaaatcggc 1920
 aactataccg gcgagatcta ccgactgctg gaagagtccc agaaccagca ggaaaagaac 1980
 gagaaggacc tgctggccct ggacagctgg aaaaatctgt ggaattgggt cgacatcacc 2040
 aactggctgt ggtacatcaa gatcttcac atgatcgtgg gggcctgat cggcctgcgg 2100
 atcatctttg ccgtgctgag catcgtgaac cgcgtgcggc agggatacag ccctctgagc 2160
 ctgcagacc tgaccagaa tccaggcggg ctggatcggc tgggccggat tgaggaagaa 2220
 ggcggcgagc aggacaagga ccgcagcatc agactcgtga acggcttctt cgctctgttt 2280
 tgggacgacc tgcggagcct gtgcctgttc tcctaccaca gactgcggga ctttatcctg 2340
 attgtggcca gagccgtcga gctgctgggc agatcttctc tgagaggcct gcagcggggc 2400
 tgggagattc tgaagtacct gggctccctg ctgcagtatt ggggcctgga actgaagaag 2460
 tccgccatca atctgctcga cacaatcgct attgccgtgg ccgaaggcac cgacagaatc 2520
 atcgagctga tccagcggat ctgccggggc atctgcaaca tccccagacg gatcagacag 2580
 ggctttgaag ccgccctcca g 2601

<210> 35

<211> 2148

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự nucleotit mã hóa cho C1D7

<400> 35

atgagagtgc ggggcattca gagaaactgg cccagtggt ggatctgggg catcctgggc 60
 ttttgatga tcattatctg ccgcgtgatg ggcaacctgt gggtcaccgt gtactacggc 120
 gtgcccgtgt ggaaagaggc caagaccacc ctgttctgcg ccagcgacgc caaggcctac 180
 gagaaagagg tgcacaacgt ctgggccacc cagcctgtg tgcccaccga cccaatccc 240
 caggaaatgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa cgacatgggtg 300
 gaccagatgc acgaggacat catccggctg tgggaccaga gcctgaagcc ctgcgtgaag 360
 ctgaccctc tgtgcgtgac cctggaatgc cggaacgtgc gcaacgtgtc cagcaacggc 420

```

acctacaata tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct      480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttttaccg gctggacatc      540
gtgcccctgg acgagaacaa cagcagcgag aagtccctccg agaacagctc cgagtactac      600
agactgatca actgcaacac cagcgccatc acccaggcct gcccgaagggt gtccttcgac      660
cctatcccca tccactactg cgcccctgcc ggctacgcca tcctgaagtg caacaacaag      720
acottcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc      780
aagcccgtgg tgtctaccca gctgctgctg aacggcagcc tggccgagga agagatcatc      840
atcagaagcg agaacctgac caacaacgcc aagacaatca tcgtccacct gaacgaaacc      900
gtgaacatca cctgtaccgg gcctaacaac aacacccgga agtccatccg gatcggccct      960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggg cactgcaat     1020
ctgagccggg acggctggaa caagacactg cagggcgctca agaagaagct ggccgaacac     1080
ttccccaaca aaaccatcaa cttcaccagc tcctctggcg gcgacctgga aatcaccacc     1140
cacagcttta actgcagagg cgagttcttc tactgcaata cctccggcct gttcaatgga     1200
acctacatgc ccaacgggac caacagcaac tccagcagca atatcaccct gccttgccgg     1260
atcaagcaga ttatcaatat gtggcaggaa gtgggcagag ctatgtacgc cctccaatc     1320
gccggcaaca tcacatgcag aagcaacatt accggcctgc tgctcaccag ggacggcggc     1380
tctaacaatg gcgtgccaaa cgacaccgag acattcagac ccggcgaggg cgacatgcgg     1440
aacaattggc ggagcgagct gtacaagtac aagtggtggg aagtgaagcc cctgggcgtg     1500
gccctaccg aagccaagag aagagtggtc gaacgcgaga agcggggccgt gggcattgga     1560
gccgtgtttc tgggaatcct gggagccgct ggagcacca tgggcgctgc ctctatcaca     1620
ctgacagtgc aggccagaca gtcctgagc ggcatcgtgc agcagcagag caacctgctg     1680
agagccatcg aggcacagca gcacatgctg cagctgaccg tgtggggcat taagcagctc     1740
cagacacggg tgctggccat tgagagatac ctgcaggatc agcagctgct cggcctgtgg     1800
ggctgtagcg gcaagctgat ctgtaccacc gccgtgcctt ggaacacctc ctggtccaac     1860
aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agaaatcggc     1920
aactataccg gcgagatcta ccgactgctg gaagagtccc agaaccagca ggaaaagaac     1980
gagaaggacc tgctggccct ggacagctgg aaaaatctgt ggaattggtt cgacatcacc     2040
aactggctgt ggtacatcaa gatcttcatc atgatcgtgg gcggcctgat cggcctgcgg     2100
atcatctttg ccgtgctgag catcgtgaac cgcgtgcggc agggctac      2148

```

<210> 36

<211> 724

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> trình tự trime Env khảm

<400> 36

Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp

1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Ala Gly Lys

20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr

35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val

50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro

65 70 75 80

Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys

85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu

115 120 125

Asn Cys Thr Asp Asp Val Arg Asn Val Thr Asn Asn Ala Thr Asn Thr

130 135 140

Asn Ser Ser Trp Gly Glu Pro Met Glu Lys Gly Glu Ile Lys Asn Cys

145 150 155 160

Ser Phe Asn Ile Thr Thr Ser Ile Arg Asn Lys Val Gln Lys Gln Tyr

165 170 175

Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asp Ser Asn

180 185 190

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

Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile
 435 440 445
 Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly
 450 455 460
 Asn Asp Thr Ser Gly Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met
 465 470 475 480
 Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile
 485 490 495
 Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Glu Arg Val Val Gln
 500 505 510
 Arg Glu Glu Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu
 515 520 525
 Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val
 530 535 540
 Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu
 545 550 555 560
 Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp
 565 570 575
 Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu
 580 585 590
 Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile
 595 600 605
 Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu
 610 615 620
 Asp Lys Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile
 625 630 635 640
 Asn Asn Tyr Thr Ser Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn
 645 650 655
 Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala
 660 665 670

Ser Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys

675

680

685

Ser Arg Ile Glu Gly Arg Gly Ser Gly Gly Tyr Ile Pro Glu Ala Pro

690

695

700

Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu

705

710

715

720

Ser Thr Phe Leu

<210> 37

<211> 28

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> Các gốc 655-682 của SEQ ID NO:18

<400> 37

Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Asn Asn Leu Trp Asn

1

5

10

15

Trp Phe Ser Ile Ser Lys Trp Leu Trp Tyr Ile Lys

20

25