



(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)



1-0030903

(51)⁷ A61K 39/12; C07K 14/005

(13) B

-
- (21) 1-2014-02136 (22) 27/11/2012
(86) PCT/EP2012/073706 27/11/2012 (87) WO/2013/079473 06/06/2013
(30) 11191003.0 28/11/2011 EP; 11191009.7 28/11/2011 EP; 61/564,198 28/11/2011 US;
61/564,086 28/11/2011 US; 12166268.8 01/05/2012 EP; 61/720,281 30/10/2012 US
(45) 25/01/2022 406 (43) 27/10/2014 319A
(73) Janssen Vaccines & Prevention B.V. (NL)
Archimedesweg 4, NL-2333 CN Leiden, The Netherlands
(72) MEIJBERG, Jan Wilem (NL); IMPAGLIAZZO, Antonietta (NL); VOGELS, Ronald
(NL); FRIESEN, Robert Heinz Edward (NL); ALARD, Philippe (BE); LOVERIX,
Stefan (BE); RADOSEVIC, Katarina (NL).
(74) Công ty Cổ phần Sở hữu công nghiệp INVESTIP (INVESTIP)
-

(54) POLYPEPTIT VÙNG GỐC HEMAGGLUTININ CỦA VIRUT CÚM VÀ CHẾ
PHẨM MIỄN DỊCH CHỨA POLYPEPTIT NÀY

(57) Sáng chế đề xuất các polypeptit vùng gốc hemagglutinin của virus cúm chứa (a) vùng HA1 của hemagglutinin của virus cúm mà bao gồm đoạn gốc đầu tận cùng đầu N của HA1, được liên kết cộng hóa trị bằng trình tự liên kết gồm 0-50 gốc axit amin với đoạn gốc đầu tận cùng đầu C của HA1, và (b) vùng HA2 của hemagglutinin của virus cúm, trong đó một hoặc nhiều axit amin trong vùng HA2 đã bị gây đột biến. Sáng chế cũng đề xuất axit nucleic mã hóa polypeptit, chế phẩm chứa polypeptit và/hoặc phân tử axit nucleic.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến lĩnh vực y học. Sáng chế đề cập đến các polypeptit vùng gốc hemagglutinin của virus cúm, phương pháp tạo ra các polypeptit vùng gốc hemagglutinin, chế phẩm chứa các polypeptit này, vaccine chứa chúng và phương pháp sử dụng chúng, cụ thể là để phát hiện, phòng ngừa và/hoặc điều trị bệnh cúm.

Tình trạng kỹ thuật của sáng chế

Virus cúm là tác nhân gây bệnh chính cho người, gây ra bệnh về đường hô hấp (thường được gọi là “bệnh cúm” hoặc “cúm”) có mức độ trầm trọng từ nhiễm cận lâm sàng đến viêm phổi cấp do virus có thể gây tử vong. Tác động lâm sàng của nhiễm cúm khác nhau tùy theo độc lực của chủng virus cúm và tình trạng phơi nhiễm, tiền sử, độ tuổi và miễn dịch của vật chủ. Mỗi năm, trên thế giới ước tính có khoảng 1 tỷ người nhiễm virus cúm, gây ra 3-5 triệu ca bệnh nghiêm trọng và ước tính có 300.000 đến 500.000 ca tử vong liên quan đến bệnh cúm. Phần lớn trong số các ca nhiễm cúm này là do virus cúm A mang các typ phụ H1 hoặc H3 hemagglutinin, và phần nhỏ hơn là do virus cúm B, và do đó đại diện của tất cả ba typ phụ này được đưa vào vaccine cúm mùa. Công tác gây miễn dịch hiện nay dựa trên sự nhận dạng sớm các virus cúm đang lưu hành để cho phép sản xuất kịp thời vaccine cúm mùa hiệu quả. Ngoài những khó khăn cố hữu trong việc dự đoán các chủng sẽ là chủng trội trong suốt mùa tới, tính kháng kháng virus và lẫn tránh miễn dịch cũng đóng vai trò trong việc làm cho vaccine hiện nay không có hiệu quả phòng ngừa tỷ lệ mắc bệnh và tỷ lệ tử vong. Ngoài ra, khả năng gây ra đại dịch do chủng virus có độc lực cao bắt nguồn từ các ổ chứa động vật và được sắp xếp lại để gia tăng sự phát tán từ người sang người, đặt ra mối đe dọa hiện thực và nghiêm trọng đến sức khỏe toàn cầu.

Virus cúm A được phân bố rộng rãi trong tự nhiên và có thể gây nhiễm cho nhiều loài chim và động vật có vú. Virus cúm là virus ARN có vỏ bọc thuộc họ Orthomyxoviridae. Bộ gen của chúng chứa tám đoạn ARN sợi đơn mã hóa cho 11 protein khác nhau, một nucleoprotein (NP), ba protein polymeraza (PA, PB1, và PB2), hai protein khung (M1 và M2), ba protein phi cấu trúc (NS1, NS2, và PB1-F2), và hai glycoprotein vỏ ngoài: hemagglutinin (HA) và neuraminidaza (NA). Các virus cúm được phân loại dựa trên sự khác nhau về cấu trúc kháng nguyên của các protein HA và NA, với các tổ hợp khác nhau của chúng đại diện các typ phụ virus cúm duy nhất và được tiếp tục phân loại thành các chủng virus cúm cụ thể. Mặc dù tất cả các typ phụ đã

biết có thể được tìm thấy ở chim, các typ phụ của virus cúm A hiện đang lưu hành ở người là H1N1 và H3N2. Phân tích phát sinh chủng loại đã chứng minh việc phân loại cấp nhỏ của hemagglutinin thành hai nhóm: *trong số các loại khác* các typ phụ H1, H2, H5 và H9 trong nhóm phát sinh chủng loại 1 và *trong số các loại khác* các typ phụ H3, H4 và H7 trong nhóm phát sinh chủng loại 2.

Chủng virus cúm typ B gây bệnh cúm lành tính ở người. Biến đổi kháng nguyên ở HA trong chủng virus cúm typ B nhỏ hơn so với biến đổi quan sát được trong chủng cúm typ A. Hai dòng khác nhau về mặt di truyền và kháng nguyên của virus cúm B đang lưu hành ở người được đại diện bởi các dòng B/Yamagata/16/88 (còn được gọi là B/Yamagata) và B/Victoria/2/87 (B/Victoria) (Ferguson et al., 2003). Mặc dù phổ bệnh do virus cúm B gây ra thường nhẹ hơn so với bệnh do virus cúm A, nhưng vẫn quan sát được bệnh trầm trọng do virus cúm B cần phải nằm viện.

Biết rằng các kháng thể trung hòa virus cúm chủ yếu là nhằm kháng lại hemagglutinin (HA). Hemagglutinin hoặc HA là glycoprotein dạng trimer được gắn vào vỏ ngoài của virus và có chức năng kép: HA chịu trách nhiệm liên kết với axit sialic thụ thể bề mặt tế bào và, sau đó hấp thu, HA làm trung gian cho quá trình dung hợp của màng virus và thể nhân dẫn đến giải phóng ARN của virus vào phần bào tan của tế bào. HA gồm vùng đầu lớn và vùng gốc nhỏ hơn. Sự gắn vào màng của virus được qua trung gian bởi trình tự gắn đầu tận cùng C được nối với vùng gốc. Protein được cắt sau dịch mã ở vòng được chỉ định tạo ra hai polypeptit, HA1 và HA2 (trình tự đầy đủ được gọi là HA0). Vùng đầu ở xa tâm màng chủ yếu có nguồn gốc từ HA1 và vùng gốc ở đầu gần màng chủ yếu từ HA2 (FIG. 1).

Lý do mà vacxin cúm mùa phải được thay đổi hàng năm là do sự biến đổi lớn của virus. Trong phân tử hemagglutinin, sự biến đổi này biểu thị cụ thể ở vùng đầu trong đó sự trôi dạt kháng nguyên và chuyển đổi kháng nguyên đã tạo ra một số lượng lớn các biến thể khác nhau. Vì đây cũng là vùng trội miễn dịch, hầu hết kháng thể trung hòa nhằm vào kháng lại vùng này và hoạt động bằng cách gây cản trở sự liên kết với thụ thể. Sự kết hợp của tính trội miễn dịch và sự biến đổi lớn của vùng đầu cũng giải thích tại sao việc nhiễm một chủng cụ thể không dẫn đến khả năng miễn dịch cho các chủng khác: các kháng thể được tạo ra do lần nhiễm đầu tiên chỉ nhận diện được số lượng chủng hạn chế có quan hệ gần với virus của lần nhiễm ban đầu.

Gần đây, các polypeptit vùng gốc hemagglutinin của virus cúm, không có toàn bộ hoặc gần như toàn bộ vùng đầu hình cầu hemagglutinin của virus cúm, đã được mô tả và được sử dụng để tạo ra đáp ứng miễn dịch với một hoặc nhiều epitop bảo thủ của polypeptit vùng gốc. Cho rằng các epitop của polypeptit vùng gốc có tính gây miễn

dịch ít hơn so với vùng có tính gây miễn dịch cao của vùng đầu hình cầu, do đó việc không có vùng đầu hình cầu trong polypeptit vùng gốc có thể cho phép đáp ứng miễn dịch kháng lại một hoặc nhiều epitop của polypeptit vùng gốc để phát triển (Steel et al., 2010). Steel và các đồng tác giả do đó đã tạo ra phân tử mới bằng cách loại bỏ gốc các axit amin từ 53 đến 276 của HA1 của chủng A/Puerto Rico/8/1934 (H1N1) và A/Hong Kong/1968 (H3N2) từ trình tự HA ban đầu, và thay thế các gốc này bằng trình tự nối ngắn linh động GGGG. Chủng ngừa cho chuột bằng cấu trúc HK68 của H3 không tạo ra kháng huyết thanh mà phản ứng chéo với các HA nhóm 1. Ngoài ra, như được thể hiện trong các ví dụ dưới đây, polypeptit vùng gốc không có tính ổn định cao và không chấp nhận cấu hình chính xác như được chứng minh bằng sự thiếu liên kết của kháng thể mà được thể hiện liên kết với epitop bảo thủ trong vùng gốc.

Ngoài ra, Bommakanti và các đồng tác giả (2010) mô tả polypeptit dựa trên HA2 bao gồm các gốc axit amin 1-172 của HA2, đoạn nối gồm 7 axit amin (GSAGSAG), các gốc axit amin 7-46 của HA1, đoạn nối gồm 6 axit amin GSAGSA, tiếp theo là các gốc 290-321 của HA1, có các đột biến V297T, I300E, Y302T và C305T trong HA1. Thiết kế này dựa trên trình tự HA của H3 (A/Hong Kong/1968). Polypeptit chỉ tạo ra sự bảo vệ chéo kháng lại một chủng virus cúm khác trong typ phụ H3 (A/Phil/2/82 nhưng không kháng lại typ phụ H1 (A/PR/8/34).

Do đó, vẫn cần có nhu cầu về vaccin phổ biến an toàn và hiệu quả, kích thích việc tạo ra đáp ứng bởi kháng thể trung hòa mạnh, phổ rộng và tạo ra sự bảo vệ kháng lại tập hợp rộng gồm các chủng virus cúm hiện nay và trong tương lai (cả cúm mùa và đại dịch), cụ thể là tạo ra sự bảo vệ kháng lại một hoặc nhiều typ phụ của virus cúm A thuộc nhóm phát sinh chủng loại 1 và/hoặc nhóm phát sinh chủng loại 2, để phòng ngừa và điều trị hiệu quả bệnh cúm.

Bản chất kỹ thuật của sáng chế

Sáng chế đề cập đến các polypeptit vùng gốc hemagglutinin của virus cúm, phương pháp tạo ra polypeptit vùng gốc, chế phẩm chứa các polypeptit này, vaccin chứa chúng và phương pháp sử dụng chúng.

Theo khía cạnh thứ nhất, sáng chế đề xuất các polypeptit gây miễn dịch mới chứa vùng gốc hemagglutinin của virus cúm và không có vùng đầu hình cầu, được gọi là polypeptit vùng gốc hemagglutinin (HA) của virus cúm. Các polypeptit này có khả năng gây đáp ứng miễn dịch khi được sử dụng cho đối tượng, cụ thể đối tượng là người. Các polypeptit theo sáng chế trình diện các epitop bảo thủ của phân tử HA vùng cuống gần gốc màng cho hệ miễn dịch khi không có các epitop trội mà có mặt trong vùng đầu ở xa màng. Cuối cùng, một phần của trình tự ban đầu của protein HA0

tạo nên vùng đầu được loại bỏ và trình tự axit amin còn lại được nối lại, trực tiếp hoặc, theo một số phương án, bằng cách đưa vào trình tự nối ngắn linh động ('cầu nối') để khôi phục tính liên tục của chuỗi axit amin. Trình tự thu được tiếp tục được cải biến bằng cách đưa vào các đột biến đặc hiệu làm ổn định cấu trúc ba chiều tự nhiên của phần còn lại của phân tử HA0. Các polypeptit gây miễn dịch không chứa HA1 và/hoặc HA2 chiều dài đủ của virus cúm.

Các polypeptit vùng gốc hemagglutinin của virus cúm này dựa trên HA của chủng virus cúm mà thường được sử dụng để sản xuất vắc xin cúm cho người. Cụ thể là, polypeptit dựa trên HA của virus cúm A thuộc các typ phụ H1, H5 và/hoặc H3.

Theo một số phương án, sáng chế đề xuất các polypeptit vùng gốc hemagglutinin của virus cúm chứa (a) vùng HA1 của hemagglutinin của virus cúm mà bao gồm đoạn gốc đầu tận cùng N của HA1, được liên kết cộng hóa trị bằng trình tự liên kết gồm 0-50 gốc axit amin với đoạn gốc đầu tận cùng C của HA1, và (b) vùng HA2 của hemagglutinin của virus cúm, trong đó polypeptit vùng gốc của hemagglutinin này kháng với sự phân cắt bằng proteaza ở vị trí nối giữa HA1 và HA2, và trong đó một hoặc nhiều axit amin trong trình tự axit amin nối chuỗi cuộn xoắn A và chuỗi cuộn xoắn CD của HA2 đã được gây đột biến so với vùng HA2 của virus cúm kiểu dại. Tốt hơn, vùng HA1 và HA2 có nguồn gốc từ virus cúm A được chọn từ nhóm bao gồm typ phụ H1, H5 và H3.

Polypeptit theo sáng chế chứa một hoặc nhiều đột biến trong trình tự axit amin của HA2 nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD, như được chỉ ra trên FIG. 1. Theo một số phương án, một hoặc nhiều axit amin kỵ nước trong trình tự axit amin của HA2 này đã được thay thế bằng các axit amin ưa nước, như các axit amin phân cực và/hoặc tích điện, hoặc axit amin linh động glyxin (G).

Theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm các axit amin 1-x của HA1, và đoạn gốc đầu tận cùng C của HA1 gồm các axit amin y-điểm cuối (tức là axit amin đầu tận cùng C của HA1) của HA1. Do đó, theo một số phương án, việc loại bỏ đoạn HA1 chứa trình tự axit amin từ axit amin ở vị trí x+1 đến và bao gồm axit amin ở vị trí y-1. Theo một số phương án, polypeptit không chứa trình tự tín hiệu. Do đó theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm các axit amin p-x của HA1, trong đó p là axit amin đầu tiên của phân tử HA trưởng thành (ví dụ p=18 trong trường hợp trình tự SEQ ID NO: 1). Người có hiểu biết trung bình trong lĩnh vực sẽ có thể tạo ra polypeptit được mô tả ở đây không có các peptit tín hiệu (ví dụ axit amin 1-17 của trình tự SEQ ID NO: 1). Theo một số phương án, polypeptit theo

sáng chế chứa trình tự nội bào của HA và vùng xuyên màng. Theo các phương án khác, polypeptit theo sáng chế không chứa trình tự nội bào của HA và vùng xuyên màng. Theo một số phương án, trình tự nội bào và xuyên màng, ví dụ trình tự axit amin từ vị trí (hoặc vị trí tương đương của) 523, 524, 525, 526, 527, 526, 528, 529, hoặc 530 của vùng HA2 đến đầu tận cùng C của vùng HA2 được loại bỏ.

Polypeptit theo sáng chế không chứa HA1 chiều dài đủ.

Theo một số phương án, các polypeptit này được glycosyl hóa.

Theo một số phương án, các polypeptit gây miễn dịch hầu như nhỏ hơn HA0, tốt hơn không có toàn bộ hoặc hầu như toàn bộ vùng đầu hình cầu của HA. Tốt hơn, các polypeptit gây miễn dịch có chiều dài không nhiều hơn 360, tốt hơn không nhiều hơn 350, 340, 330, 320, 310, 305, 300, 295, 290, 285, 280, 275, hoặc 270 axit amin. Theo một số phương án, các polypeptit gây miễn dịch có chiều dài trong khoảng từ 250 đến 350, tốt hơn là từ 260 đến 340, tốt hơn là từ 270 đến 330, tốt hơn là từ 270 đến 330 axit amin.

Theo một số phương án, polypeptit còn chứa một hoặc nhiều đột biến khác trong vùng HA1 và/hoặc HA2, khi so với trình tự axit amin của HA mà các vùng HA1 và HA2 dựa vào.

Sáng chế còn đề cập đến phương pháp tạo ra các polypeptit vùng gốc hemagglutinin của virus cúm, bao gồm các bước phổ biến sau:

- (a) Cung cấp trình tự axit amin HA0 của virus cúm;
- (b) Loại bỏ vị trí cắt giữa HA1 và HA2;
- (c) Loại bỏ trình tự axit amin của vùng đầu hình cầu từ trình tự HA0, cụ thể là trình tự axit amin bắt đầu từ vị trí x+1 đến y-1;
- (d) Đưa một hoặc nhiều đột biến vào trong trình tự axit amin nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD; và
- (e) Đưa một hoặc nhiều cầu disulfid vào polypeptit vùng gốc của HA.

Các polypeptit có thể thu được bằng phương pháp này cũng là một phần của sáng chế.

Theo một số phương án polypeptit chứa các epitop vùng gốc bảo thủ của nhóm 1 trung hòa chéo kháng thể CR6261 (như được bộc lộ trong WO2008/028946) và/hoặc của kháng thể CR9114 (như được mô tả dưới đây và trong đơn đang được theo đuổi đồng thời EP 11173953.8), kháng thể có khả năng liên kết với và trung hòa cả hai virus cúm A thuộc nhóm 1 và nhóm 2, cũng như virus cúm B. Do đó, một khía cạnh khác của sáng chế đề xuất polypeptit vùng gốc của HA, trong đó polypeptit này liên kết với kháng thể CR6261 và/hoặc kháng thể CR9114. Theo một phương án,

polypeptit không liên kết với CR8057 (được mô tả trong WO 2010/130636), kháng thể đơn dòng chỉ liên kết với virus cúm H3. Theo một số phương án, polypeptit liên kết với kháng thể CR8020, CR8043 và/hoặc CR9114. Các polypeptit vùng gốc hemagglutinin của virus cúm được đề cập trong sáng chế thích hợp để sử dụng trong chế phẩm gây miễn dịch (ví dụ vaccin) có khả năng tạo ra các đáp ứng miễn dịch kháng lại nhiều chủng virus cúm A và/hoặc B. Theo một phương án, các polypeptit vùng gốc hemagglutinin của virus cúm này có khả năng tạo ra các đáp ứng miễn dịch kháng lại chủng virus cúm A thuộc nhóm phát sinh chủng loại 1 và/hoặc 2, cụ thể là kháng lại các chủng virus cúm thuộc cả hai nhóm phát sinh chủng loại 1 và 2. Theo một phương án, polypeptit có khả năng tạo ra đáp ứng miễn dịch kháng lại các chủng virus cúm cùng loài. Theo một phương án, polypeptit có khả năng tạo ra đáp ứng miễn dịch kháng lại các chủng virus cúm khác loài của cùng và/hoặc các typ phụ khác nhau. Theo phương án khác, polypeptit có khả năng tạo ra đáp ứng miễn dịch đối với các chủng virus cúm thuộc cả hai nhóm phát sinh chủng loại 1 và 2 và các chủng virus cúm B.

Polypeptit theo sáng chế có thể được sử dụng ví dụ trong một liệu pháp điều trị và/hoặc dự phòng và/hoặc chẩn đoán tiêu chuẩn bệnh hoặc tình trạng bệnh lý do virus cúm gây ra, cụ thể là virus cúm A thuộc nhóm phát sinh chủng loại 1 hoặc 2 và/hoặc virus cúm B, hoặc kết hợp với các phép điều trị dự phòng và/hoặc chữa bệnh khác, như vaccin, tác nhân kháng virus và/hoặc kháng thể đơn dòng (đã có hoặc trong tương lai).

Theo khía cạnh khác, sáng chế đề xuất các phân tử axit nucleic mã hóa polypeptit vùng gốc của HA của virus cúm. Theo khía cạnh khác nữa, sáng chế đề xuất các vector chứa axit nucleic mã hóa các polypeptit gây miễn dịch.

Theo khía cạnh khác, sáng chế đề xuất các phương pháp gây đáp ứng miễn dịch ở đối tượng, phương pháp này bao gồm việc cho đối tượng sử dụng polypeptit và/hoặc phân tử axit nucleic theo sáng chế.

Theo khía cạnh khác, sáng chế đề xuất các chế phẩm gây miễn dịch chứa polypeptit và/hoặc phân tử axit nucleic theo sáng chế. Các chế phẩm gây miễn dịch được đề xuất theo sáng chế có thể ở dạng bất kỳ mà cho phép chế phẩm được sử dụng cho đối tượng, ví dụ chuột, chồn hoặc người. Theo phương án cụ thể, các chế phẩm gây miễn dịch là thích hợp để sử dụng cho người. Polypeptit, các phân tử axit nucleic và chế phẩm có thể được sử dụng trong các phương pháp phòng ngừa và/hoặc điều trị bệnh do virus cúm và/hoặc cho mục đích chẩn đoán bệnh. Các chế phẩm này có thể còn chứa chất mang được dụng hoặc tá dược. Theo một số phương án, các chế phẩm được mô tả ở đây bao gồm, hoặc được sử dụng kết hợp với, tá dược.

Theo khía cạnh khác, sáng chế đề xuất các polypeptit, axit nucleic và/hoặc các chế phẩm gây miễn dịch được sử dụng làm vaccin. Cụ thể, sáng chế đề cập đến các polypeptit gây miễn dịch, axit nucleic, và/hoặc các chế phẩm gây miễn dịch được sử dụng làm vaccin theo sáng chế và/hoặc điều trị bệnh hoặc tình trạng bệnh lý do virus cúm gây ra A thuộc typ phụ của nhóm phát sinh chủng loại 1 và/hoặc 2 và/hoặc virus cúm B.

Các phương án khác và việc sử dụng polypeptit theo sáng chế sẽ trở nên rõ ràng hơn khi tham khảo phần mô tả chi tiết sáng chế sau đây.

Mô tả vắn tắt các hình vẽ

FIG. 1: Thể hiện mô hình của monome HA ở trạng thái trước khi dung hợp khi biểu hiện ở dạng trime tự nhiên. HA1 được thể hiện bằng màu xám nhạt, HA2 được thể hiện bằng màu xám đậm. Chuỗi cuộn xoắn A (một phần quan trọng của epitop của CR6261) và chuỗi cuộn xoắn CD (một phần của bề mặt chung của trime) được chỉ rõ, là vòng nối các yếu tố cấu trúc bậc hai này.

FIG. 2: Liên kết của kháng thể đơn dòng với HA chiều dài đủ và polypeptit vùng gốc của HA theo sáng chế như được phân tích bởi FACS. A: Tỷ lệ phần trăm tế bào dương tính sau khi nhuộm. B: cường độ phát huỳnh quang trung bình. H1-Chiều dài đủ (SEQ ID NO: 1), miniHA-cl1 (SEQ ID NO: 3), miniHA-cl1+2 (SEQ ID NO: 4), miniHA-cl1+3 (SEQ ID NO: 5), miniHA-cl1+4 (SEQ ID NO: 6) miniHA-cl1+2+3 (SEQ ID NO: 7), miniHA-cl1+2+3+4 (SEQ ID NO: 8).

FIG. 3: Liên kết của kháng thể đơn dòng với HA chiều dài đủ và polypeptit vùng gốc của HA như được phân tích bởi FACS. A: Tỷ lệ phần trăm tế bào dương tính sau khi nhuộm. B: cường độ phát huỳnh quang trung bình. H1-Chiều dài đủ (SEQ ID NO: 1), miniHA (SEQ ID NO: 2), miniHA-cl1 (SEQ ID NO: 3).

FIG. 4: Liên kết của các kháng thể trong huyết thanh với HEK293F biểu hiện HA chiều dài đủ và polypeptit theo sáng chế. A: cường độ phát huỳnh quang trung bình. B: Tỷ lệ phần trăm tế bào dương tính sau khi nhuộm. H1-FL (SEQ ID NO: 1), CL1 (SEQ ID NO: 3), CL1+2 (SEQ ID NO: 4) và CL1+4 (SEQ ID NO: 6). cM2 là đối chứng âm.

FIG. 5: Liên kết của kháng thể đơn dòng với HA chiều dài đủ và polypeptit vùng gốc của HA như được phân tích bởi FACS. Đỉnh: Tỷ lệ phần trăm tế bào dương tính sau khi nhuộm. Đáy: cường độ phát huỳnh quang trung bình. H1-Chiều dài đủ (SEQ ID NO: 1), miniHA-cl1 (SEQ ID NO: 3), H1-mini1-cl11 (SEQ ID NO: 9), H1-mini2-cl11 (SEQ ID NO: 10), H1-mini3-cl11 (SEQ ID NO: 11), H1-mini4-cl11 (SEQ

ID NO: 12), H1-mini1-cl11+5 (SEQ ID NO: 13), H1-mini2-cl11+5 (SEQ ID NO: 14), H1 mini3-cl11+5 (SEQ ID NO: 15), và H1-mini4-cl11+5 (SEQ ID NO: 16).

FIG. 6: Liên kết của các kháng thể trong huyết thanh với vùng ngoài của HA chiều dài đủ từ A/Brisbane 59/2007 sau khi gây miễn dịch bằng cách tiêm bắp ADN mã hóa HA của A/Brisbane/59/2007 (SEQ ID NO: 1), miniHA-cụm1 (SEQ ID NO: 3), Mini2-cụm11 (SEQ ID NO: 10), Mini1-cụm11+5 (SEQ ID NO: 13), Mini2-cụm11+5 (SEQ ID NO: 14) và cM2 (trình tự M2 liên ứng) hoặc gây miễn dịch bằng súng bằng gen ADN mã hóa HA của A/Brisbane/59/2007 (SEQ ID NO: 1), Mini2-cụm11+5 (SEQ ID NO: 14) và cM2 (trình tự M2 liên ứng). Dải A: 28 ngày sau khi gây miễn dịch lần đầu. Dải B: sau 49 ngày gây miễn dịch.

FIG. 7: Liên kết của kháng thể đơn dòng với HA chiều dài đủ và polypeptit vùng gốc của HA như được phân tích bởi FACS. Định: Tỷ lệ phần trăm tế bào dương tính sau khi nhuộm. Đáy: cường độ phát huỳnh quang trung bình. H1-Chiều dài đủ (SEQ ID NO: 1), miniHA (SEQ ID NO: 2), H1-mini2-cl11+5 (SEQ ID NO: 14), H1-mini2-cl1 +5 (SEQ ID NO: 48), H1-mini2-cl1+5+6 (SEQ ID NO: 46), H1-mini2-cl11+5+6 (SEQ ID NO: 47), H1-mini2-cl1+5+6-trim (SEQ ID NO: 44), H1-mini2-cl1+5+6-GCN4 (SEQ ID NO: 45).

FIG. 8: Liên kết của kháng thể đơn dòng với HA chiều dài đủ và polypeptit vùng gốc của HA như được phân tích bởi FACS. A: Tỷ lệ phần trăm tế bào dương tính sau khi nhuộm. B: cường độ phát huỳnh quang trung bình.

FIG. 9: Liên kết của kháng thể đơn dòng với HA chiều dài đủ và polypeptit vùng gốc của HA như được phân tích bởi FACS. A: Tỷ lệ phần trăm tế bào dương tính sau khi nhuộm. B: cường độ phát huỳnh quang trung bình.

FIG. 10: Sự biểu hiện của các cấu trúc dựa trên chủng Hong Kong/1/1968 trên bề mặt tế bào.

FIG. 11: Phân tích SDS-PAGE (A-D) và thẩm tách miễn dịch - Western Blot (E-F) về việc tinh chế một số polypeptit theo sáng chế. Đối với Western Blot kháng thể nhằm kháng lại đuôi his được sử dụng để phát hiện.

FIG. 12: Liên kết của kháng thể đơn dòng CR9114 (A), CR8020 (B) và kháng thể đa dòng kháng huyết thanh chứa HA của H1 (C) với một số polypeptit theo sáng chế như được phát hiện bằng ELISA.

FIG. 13: Phân tích SDS-PAGE (A) và Western Blot (B) về sự glycosyl hóa của polypeptit theo sáng chế. Khi khử glycosyl hóa, các dải khuếch tán tập trung lại ở khối lượng phân tử mong muốn. Đối với Western Blot huyết thanh đa dòng kháng lại H1 HA được sử dụng để phát hiện.

FIG. 14: Phân tích SEC-MALS của polypeptit theo sáng chế. Các vết được đánh dấu bằng trình tự SEQ ID NO.

FIG 15: A: Phân tích Western Blot dịch nổi của tế bào biểu hiện SEQ ID NO: 145. Đối với Western Blot, kháng thể kháng đuôi his được sử dụng để phát hiện B: Liên kết của kháng thể đơn dòng CR9114 (hình vuông), CR6261(hình tròn), CR8020 (hình tam giác ngược) và FI6v3 (hình tam giác xuôi) với trình tự SEQ ID NO: 145 như được phát hiện bằng ELISA.

FIG. 16: Thông số rửa giải của quá trình tinh sạch trình tự SEQ ID NO: 145 từ dịch nổi nuôi cấy trên cột bắt giữ His. Polypeptit theo sáng chế rửa giải ở 100 (đỉnh A) và 200 mM (đỉnh B) imidazol, tương ứng.

FIG. 17: A-B: Thông số rửa giải của quá trình tinh sạch trình tự SEQ ID NO: 145 bằng sắc ký theo loại cỡ (Superdex 200). Cả hai đỉnh A và B chứa polypeptit theo sáng chế. C: Phân tích PAGE tự nhiên các phân đoạn từ sắc ký theo loại cỡ. Phần lớn protein được tinh sạch chạy ở khối lượng phân tử đồng nhất với dạng monome của protein. D: Phân tích SDS PAGE các phân đoạn từ sắc ký theo loại cỡ.

FIG. 18: Tiến trình thời gian của đáp ứng IgG hướng đến vùng ngoài của protein cùng loại chiều dài đủ theo kết quả tiến trình gây miễn dịch bằng ADN được mô tả trong đơn yêu cầu cấp patent này.

FIG. 19: Đáp ứng IgG ở tuần 7 sau lần gây miễn dịch ban đầu cho cá thể chuột kháng vùng ngoài của hemagglutinin chiều dài đủ từ chủng cùng loài H1N1 A/Brisbane/59/2007 (A) và chủng khác loài H1N1 A/California/07/2009 (B). Các ký hiệu mở tương ứng với các giá trị thấp hơn giới hạn phát hiện của thử nghiệm.

FIG. 20: Đáp ứng IgG ở tuần 7 sau lần gây miễn dịch ban đầu cho cá thể chuột kháng vùng ngoài của hemagglutinin chiều dài đủ từ chủng cùng loài H1N1 A/Brisbane/59/2007 (A), chủng khác loài H1N1 A/California/07/2009 (B) chủng khác typ phụ H5N1 A/Vietnam/1203/2004 (C) và chủng khác typ phụ H3N2 A/Hong Kong/1/1968 (D). Các ký hiệu mở tương ứng với các giá trị thấp hơn giới hạn phát hiện của thử nghiệm.

FIG. 21: Thử nghiệm FACS polypeptit vùng gốc dựa trên HA của H3. Cường độ phát huỳnh quang trung bình (A) và % tế bào dương tính (B) được thể hiện.

FIG. 22: Đáp ứng IgG ở tuần 7 sau lần gây miễn dịch ban đầu cho cá thể chuột kháng hemagglutinin chiều dài đủ từ chủng cùng loài H1N1 A/Brisbane/59/2007 (dải A), chủng khác loài H1N1 A/California/07/2009 (dải B) và chủng khác typ phụ H5N1 A/Vietnam/1203/2004 (dải C). Các ký hiệu mở tương ứng với các giá trị thấp hơn giới hạn phát hiện của thử nghiệm.

FIG. 23: Thử nghiệm FACS liên kết của các mAb CR6261, CR9114, CR8020 và CR9020, cũng như huyết thanh kháng thể đa dòng kháng H1 với HA chiều dài đủ và polypeptit tương ứng theo sáng chế. Đỉnh: Cường độ phát huỳnh quang trung bình. Đáy: tỷ lệ phần trăm tế bào dương tính. Các thanh đặc biểu thị protein chiều dài đủ, các thanh sọc biểu thị polypeptit theo sáng chế. HA chiều dài đủ và the polypeptit tương ứng theo sáng chế có nguồn gốc từ trình tự có cùng màu nền.

FIG. 24: Đường cong sống sót Kaplan-Meier (A), thay đổi về khối lượng (B) và điểm lâm sàng trung bình (C) đối với thử nghiệm thử thách virus cúm được mô tả trong ví dụ 21.

FIG. 25: So sánh trình tự H1N1 được chọn theo ví dụ 22.

FIG. 26: Thử nghiệm FACS polypeptit vùng gốc dựa trên HA của H1 được chọn theo ví dụ 22. Cường độ phát huỳnh quang trung bình được thể hiện.

FIG. 27: Động học của liên kết của HA chiều dài đủ của H1 (SEQ ID NO 149) ở dạng trime và monome của nó và s-H1-mini2-cụm1+5+6-GCN4(SEQ ID NO: 145) với kháng thể đơn dòng được cố định CR6261 (A), CR9114 (B) và CR8020 (C) như được xác định bằng phương pháp đo giao thoa lớp sinh học.

FIG. 28: Chuẩn độ ở trạng thái ổn định liên kết của s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) với CR6261 (A) và CR9114 (B) được cố định sau đó đo giao thoa lớp sinh học.

FIG. 29: Đáp ứng IgG kháng vùng ngoài của HA từ chủng A/Hong Kong/1/1968 ở ngày 49 như được xác định bằng ELISA. Chi tiết thử nghiệm được mô tả trong ví dụ 24. Các ký hiệu mở tương ứng với các giá trị thấp hơn giới hạn phát hiện của thử nghiệm.

FIG. 30: Đường cong sống sót Kaplan-Meier (A), thay đổi về khối lượng (B) và điểm lâm sàng trung bình (C) đối với thử nghiệm thử thách virus cúm được mô tả trong ví dụ 24.

FIG. 31: Thử nghiệm FACS polypeptit vùng gốc dựa trên HA của H1 được chọn theo ví dụ 25. Cường độ phát huỳnh quang trung bình được thể hiện.

FIG. 32: Đáp ứng IgG sau 49 ngày kháng vùng ngoài của HA như được xác định bằng HA từ chủng A/Wisconsin/67/2005 (A), A/Hong Kong/1/1968 (B) và A/Perth/16/2009 (C). Chi tiết thử nghiệm được mô tả trong ví dụ 26. Các ký hiệu mở tương ứng với các giá trị thấp hơn giới hạn phát hiện của thử nghiệm.

FIG. 33: Đáp ứng IgG sau 49 ngày kháng vùng ngoài của HA từ A/Hong Kong/1/1968 như được xác định bằng ELISAs. Chi tiết thử nghiệm được mô tả trong

ví dụ 27. Các ký hiệu mở tương ứng với các giá trị thấp hơn giới hạn phát hiện của thử nghiệm.

FIG. 34: Thử nghiệm FACS polypeptit vùng gốc dựa trên HA của H1 được chọn theo ví dụ 28. Cường độ phát huỳnh quang trung bình được thể hiện.

Mô tả chi tiết sáng chế

Định nghĩa các thuật ngữ được sử dụng trong bản mô tả được đưa ra như sau.

Axit amin theo sáng chế có thể là axit amin bất kỳ trong số 20 axit amin có trong tự nhiên (hoặc axit amin 'chuẩn') hoặc biến thể của chúng, ví dụ như D-prolin (chất đồng phân đối ảnh dạng D của prolin), hoặc biến thể bất kỳ không được tìm thấy trong tự nhiên trong các protein, ví dụ như norleuxin. Axit amin chuẩn có thể được chia thành một số nhóm dựa trên các đặc tính của chúng. Các yếu tố quan trọng là đặc tính tích điện, ưa nước hoặc kỵ nước, các nhóm chức và kích thước. Các đặc tính này là quan trọng đối với cấu trúc protein và tương tác protein-protein. Một số axit amin có các đặc tính đặc biệt như xystein, có thể tạo thành liên kết disulfit cộng hóa trị (hoặc cầu disulfit) với các gốc xystein khác, prolin có thể tạo vòng bộ khung polypeptit, và glyxin linh động hơn các axit amin khác. Bảng 5 thể hiện các chữ viết tắt và đặc tính của các axit amin chuẩn.

Thuật ngữ "độ tương đồng về trình tự axit amin" chỉ mức độ tương đồng hoặc giống nhau giữa một cặp trình tự axit amin được so sánh, thường được biểu hiện dưới dạng phần trăm. Tỷ lệ phần trăm tương đồng là tỷ lệ phần trăm của các gốc axit amin trong trình tự ứng cử mà tương đồng hoặc giống nhau (tức là, các gốc axit amin ở vị trí đưa ra trong việc sắp xếp và so sánh là cùng axit amin) hoặc tương tự (tức là, thay thế axit amin ở vị trí đưa ra trong việc sắp xếp và so sánh là sự thay thế bảo thủ, như được thảo luận dưới đây), với gốc axit amin tương ứng trong peptit sau khi sắp xếp và so sánh các trình tự này và đưa vào các khoảng trống, nếu cần, để đạt được mức độ tương đồng về trình tự lớn nhất. Tính tương đồng về trình tự, bao gồm tỷ lệ phần trăm của trình tự tương đồng và trình tự tương tự, được xác định bằng các kỹ thuật so sánh trình tự đã biết trong lĩnh vực, như bằng cách kiểm tra bằng mắt thường và tính toán toán học, hoặc tốt hơn, việc so sánh được thực hiện bằng cách so sánh thông tin về trình tự nhờ sử dụng chương trình máy tính. Ví dụ, chương trình máy tính được ưu tiên là Genetics Computer Group (GCG; Madison, Wis.) Wisconsin package version 10.0 program, 'GAP' (Devereux et al. (1984)).

"Sự thay thế bảo thủ" để chỉ sự thay thế một axit amin của một nhóm bằng một axit amin khác của cùng nhóm đó. Theo các phương án cụ thể, sự thay thế bảo thủ không làm thay đổi cấu trúc hoặc chức năng, hoặc cả cấu trúc và chức năng của

polypeptit. Các nhóm axit amin cho mục đích thay thế bảo thủ bao gồm nhóm kỵ nước (ví dụ Met, Ala, Val, Leu), nhóm ưa nước trung tính (ví dụ Cys, Ser, Thr), nhóm axit (ví dụ Asp, Glu), nhóm bazơ (ví dụ Asn, Gln, His, Lys, Arg), nhóm phá vỡ cấu hình (ví dụ Gly, Pro) và nhóm thơm (ví dụ Trp, Tyr, Phe).

Như được sử dụng ở đây, các thuật ngữ "bệnh" và "rối loạn" được sử dụng thay thế nhau để chỉ tình trạng bệnh lý ở đối tượng. Theo một số phương án, tình trạng bệnh lý này là nhiễm virus, cụ thể là nhiễm virus cúm. Theo các phương án cụ thể, thuật ngữ "bệnh" chỉ trạng thái bệnh lý có nguyên nhân từ sự có mặt của virus trong tế bào hoặc đối tượng, hoặc do sự xâm lấn của virus vào tế bào hoặc đối tượng. Theo một số phương án, tình trạng bệnh lý là bệnh ở đối tượng, mức độ nghiêm trọng của tình trạng bệnh lý giảm do gây đáp ứng miễn dịch ở đối tượng thông qua việc sử dụng chế phẩm gây miễn dịch.

Như được sử dụng ở đây, thuật ngữ "lượng có hiệu quả" trong việc sử dụng chất điều trị bệnh cho đối tượng để chỉ lượng chất điều trị bệnh có (các) hiệu quả phòng và/hoặc điều trị bệnh. Theo một số phương án, "lượng có hiệu quả" trong việc sử dụng chất điều trị bệnh cho đối tượng chỉ lượng chất điều trị đủ để đạt được sự giảm hoặc làm thuyên giảm mức độ trầm trọng của nhiễm virus cúm, bệnh hoặc rối loạn đi kèm với nhiễm virus cúm, như, nhưng không chỉ giới hạn ở giảm thời gian kéo dài của nhiễm virus cúm, bệnh hoặc rối loạn đi kèm với nhiễm virus cúm, phòng ngừa sự tiến triển của nhiễm virus cúm, bệnh hoặc rối loạn đi kèm với nhiễm virus cúm, phòng ngừa sự phát triển, khởi phát hoặc tái diễn của nhiễm virus cúm, bệnh hoặc rối loạn đi kèm với nhiễm virus cúm, phòng ngừa hoặc làm giảm sự phát tán của virus cúm đối tượng này sang đối tượng khác, làm giảm việc phải nằm viện cho đối tượng và/hoặc thời gian nằm viện, tăng khả năng sống sót của đối tượng nhiễm virus cúm hoặc bệnh đi kèm với nhiễm virus cúm, loại bỏ nhiễm virus cúm hoặc bệnh đi kèm với nhiễm virus cúm, ức chế hoặc làm giảm sự sao chép của virus cúm, làm giảm độ chuẩn của virus cúm; và/hoặc tăng cường và/hoặc cải thiện các tác dụng phòng hoặc chữa bệnh của chất điều trị bệnh khác. Theo một số phương án, lượng có hiệu quả không mang lại sự bảo vệ hoàn toàn trước bệnh nhiễm virus cúm nhưng có tác dụng làm giảm độ chuẩn hoặc giảm số lượng virus cúm so với đối tượng không được điều trị. Lợi ích của việc làm giảm độ chuẩn, số lượng hoặc tổng gánh nặng của virus cúm bao gồm, nhưng không chỉ giới hạn ở, làm giảm các triệu chứng nghiêm trọng của nhiễm virus cúm, làm giảm các triệu chứng của nhiễm virus cúm và giảm thời gian của bệnh đi kèm với nhiễm virus cúm.

Thuật ngữ “vật chủ”, như được sử dụng ở đây, để chỉ sinh vật hoặc tế bào mà vector như vector tách dòng hoặc vector biểu hiện được đưa vào. Sinh vật hoặc tế bào có thể là nhân sơ hoặc nhân thật. Tốt hơn, vật chủ bao gồm tế bào chủ được phân lập, ví dụ tế bào chủ trong nuôi cấy. Thuật ngữ “tế bào chủ” chỉ có nghĩa là tế bào được cải biến đối với (trên)-sự biểu hiện của polypeptit theo sáng chế. Cần phải hiểu rằng thuật ngữ vật chủ không chỉ nhằm đề cập đến đối tượng là sinh vật hoặc tế bào cụ thể mà còn bao gồm thể hệ con của sinh vật hoặc tế bào này. Bởi vì một số cải biến có thể xảy ra trong các thể hệ kế tiếp do đột biến hoặc ảnh hưởng của môi trường, thực tế thể hệ con này có thể không giống hệt sinh vật hoặc tế bào cha mẹ nhưng vẫn được bao gồm trong phạm vi của thuật ngữ “vật chủ” như được sử dụng ở đây.

Thuật ngữ “được bao gồm” hoặc “bao gồm” như được sử dụng ở đây được xem là được theo sau bởi cụm từ “không giới hạn”.

Như được sử dụng ở đây, thuật ngữ “nhiễm” chỉ sự xâm lấn bởi, sự nhân lên và/hoặc sự có mặt của virus trong tế bào hoặc đối tượng. Theo một phương án, sự nhiễm là sự nhiễm “hoạt động”, tức là, sự nhiễm mà trong đó virus đang sao chép trong tế bào hoặc đối tượng. Sự nhiễm này được đặc trưng bởi sự phát tán của virus sang tế bào, mô và/hoặc cơ quan khác, từ từ tế bào, mô và/hoặc cơ quan bị nhiễm virus ban đầu. Sự nhiễm cũng có thể là sự nhiễm tiềm ẩn, tức là, sự nhiễm mà trong đó virus không sao chép. Theo một số phương án, sự nhiễm chỉ trạng thái bệnh lý bắt nguồn từ sự có mặt của virus trong tế bào hoặc đối tượng, hoặc do sự xâm lấn của virus vào tế bào hoặc đối tượng.

Virus cúm được phân thành các loại virus cúm: chi A, B và C. Thuật ngữ “typ phụ của virus cúm” như được sử dụng ở đây chỉ các biến thể của virus cúm A được đặc trưng bởi sự kết hợp của các protein bề mặt của virus cúm là hemagglutinin (H) và neuramidaza (N). Theo sáng chế các typ phụ của virus cúm có thể được thể hiện bằng số H của chúng, ví dụ như “virus cúm chứa HA của typ phụ H3”, “virus cúm thuộc typ phụ H3” hoặc “virus cúm H3”, hoặc bằng sự kết hợp của số H và số N, ví dụ như “typ phụ của virus cúm H3N2” hoặc “H3N2”. Thuật ngữ “typ phụ” bao gồm cụ thể tất cả các chủng đơn, thuộc mỗi typ phụ, thường là kết quả của các đột biến và thể hiện các thống số gây bệnh khác nhau, bao gồm các chủng phân lập trong tự nhiên cũng như các thể đột biến nhân tạo hoặc “thể tái sắp xếp và thể tương tự. Các chủng này cũng có thể được đề cập dưới dạng các “thể phân lập” khác nhau của typ phụ virus cúm. Do đó, như được sử dụng ở đây, các thuật ngữ “chủng” và “thể phân lập” có thể được sử dụng thay đổi cho nhau. Danh pháp hiện nay đối với các chủng virus cúm hoặc các thể phân lập được từ người bao gồm typ (chi) virus, tức là A, B hoặc C, vị trí địa lý lần đầu tiên

phân lập được, số chủng và năm phân lập được, thường với việc mô tả kháng nguyên của HA và NA được đưa ra trong ngoặc đơn, ví dụ A/Moscow/10/00 (H3N2). Các chủng không gây bệnh cho người cũng bao gồm vật chủ có nguồn gốc trong danh pháp này. Typ phụ của virus cúm A có thể còn được phân loại bằng cách việ dẫn đến nhóm phát sinh chủng loại của chúng. Phân tích phát sinh chủng loại đã chứng minh việc chia nhỏ các hemagglutinin thành hai nhóm chính: *trong số các loại khác* các typ phụ H1, H2, H5 và H9 trong nhóm phát sinh chủng loại 1 (virus cúm “nhóm 1”) và *trong số các loại khác* các typ phụ H3, H4, H7 và H10 trong nhóm phát sinh chủng loại 2 (virus cúm “nhóm 2”).

Như được sử dụng ở đây, thuật ngữ "bệnh do virus cúm" chỉ trạng thái bệnh lý bắt nguồn từ sự có mặt của virus cúm, ví dụ virus cúm A hoặc B trong tế bào hoặc đối tượng hoặc sự xâm lấn của virus cúm vào tế bào hoặc đối tượng. Theo các phương án cụ thể, thuật ngữ này chỉ bệnh ở đường hô hấp do virus cúm gây ra.

Như được sử dụng ở đây, thuật ngữ "axit nucleic" được dự định bao gồm các phân tử ADN (ví dụ, cADN hoặc ADN bộ gen) và các phân tử ARN (ví dụ, mARN) và các phân tử tương tự ADN hoặc ARN được tạo ra bằng cách sử dụng các phân tử tương tự axit nucleic. Axit nucleic có thể là sợi đơn hoặc sợi đôi. Các phân tử axit nucleic có thể được cải biến về mặt hóa học hoặc hóa sinh học hoặc có thể chứa các bazơ nucleotit không có trong tự nhiên hoặc dẫn xuất bazơ nucleotit, như được hiểu dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực. Các cải biến này bao gồm, ví dụ, đánh dấu, metyl hóa, thay thế một hoặc nhiều nucleotit có trong tự nhiên bằng chất tương tự, cải biến bên trong nucleotit như các liên kết không tích điện (ví dụ, metyl phosphonat, phosphotrieste, phosphoramidat, carbamat, etc.), các liên kết tích điện (ví dụ, phosphorothioat, phosphorodithioat, etc.), các gốc nhô ra (ví dụ, các polypeptit), các chất xen vào (ví dụ, acridin, psoralen, etc.), các chất cang hóa, các chất alkyl hóa, và các liên kết được cải biến (ví dụ, axit nucleic alpha anomeric, v.v.). Đề cập đến trình tự axit nucleic bao hàm trình tự bổ sung của nó trừ khi được qui định khác. Do đó, đề cập đến phân tử axit nucleic có trình tự cụ thể cần được hiểu là bao hàm sợi bổ sung của nó, với trình tự bổ sung của nó. Các sợi bổ sung cũng hữu dụng, ví dụ, trong liệu pháp điều trị đối nghĩa, các mẫu dò lai hóa và các môi trường PCR.

Như được sử dụng ở đây, theo một số phương án số lượng axit amin trong HA là dựa trên số lượng axit amin trong HA0 virus cúm kiểu đại, ví dụ số lượng axit amin của chủng virus cúm H1N1 A/Brisbane/59/2007 (SEQ ID NO: 1). Như được sử dụng trong sáng chế, cụm từ “axit amin ở vị trí “x” trong HA” do đó chỉ axit amin tương ứng với axit amin ở vị trí x trong HA0 virus cúm kiểu đại cụ thể, ví dụ

A/Brisbane/59/2007 (SEQ ID NO: 1; trong đó axit amin của vùng HA2 được thể hiện bằng chữ in nghiêng). Người có hiểu biết trung bình trong lĩnh vực cần phải hiểu rằng axit amin tương đương trong các chủng virus cúm và/hoặc typ phụ khác có thể được xác định bằng cách sắp xếp so sánh nhiều trình tự (xem ví dụ Bảng 8). Lưu ý rằng, trong hệ thống đánh số được sử dụng trong toàn bộ đơn yêu cầu cấp patent này 1 chỉ axit amin đầu tận cùng N của protein HA0 chưa trưởng thành (SEQ ID NO: 1). Trình tự trưởng thành bắt đầu ví dụ ở vị trí 18 của trình tự SEQ ID NO: 1. Theo một số phương án, số lượng axit amin tương đương là dựa trên số lượng axit amin trong HA của H30, cụ thể là số lượng axit amin chủng virus cúm H3N2 A/Wisconsin/67/2005 (SEQ ID NO: 89). Axit amin tương đương trong các trình tự HA khác của H3 có thể được xác định bằng cách sắp xếp, so sánh trình tự. Người có hiểu biết trung bình trong lĩnh vực cần phải hiểu rằng trình tự dẫn đầu (hoặc trình tự tín hiệu) mà dẫn hướng việc vận chuyển của protein trong quá trình sản sinh (ví dụ tương ứng với axit amin 1-17 của trình tự SEQ ID NO: 89), thường không có mặt trong polypeptit cuối cùng, ví dụ được sử dụng trong vacxin. Theo một số phương án, polypeptit theo sáng chế do đó bao gồm trình tự axit amin không có trình tự dẫn đầu, tức là trình tự axit amin này dựa trên trình tự axit amin của HA0 không có trình tự tín hiệu.

"Polypeptit" chỉ polyme gồm các axit amin được liên kết với nhau bằng liên kết amit mà người có hiểu biết trung bình trong lĩnh vực đã biết. Như được sử dụng ở đây, thuật ngữ này có thể chỉ một chuỗi polypeptit được liên kết bằng các liên kết amit cộng hóa trị. Thuật ngữ này cũng có thể chỉ nhiều chuỗi polypeptit được liên kết bằng các mối tương tác không cộng hóa trị như liên kết ion, liên kết hydro, liên kết Van der Waals và liên kết kỵ nước. Người có hiểu biết trung bình trong lĩnh vực sẽ nhận ra rằng thuật ngữ này bao gồm các polypeptit được cải biến, ví dụ bằng quá trình xử lý sau dịch mã như cắt peptit tín hiệu, tạo liên kết disulfid, glycosyl hóa (ví dụ, glycosyl hóa được liên kết N), cắt bằng proteaza và cải biến lipid (ví dụ S-palmitoyl hóa).

"Polypeptit vùng gốc" chỉ polypeptit mà bao gồm một hoặc nhiều chuỗi polypeptit tạo ra vùng gốc của hemagglutinin (HA) có trong tự nhiên (hoặc kiểu đại). Thông thường, polypeptit vùng gốc là một chuỗi polypeptit (tức là tương ứng với vùng gốc của polypeptit hemagglutinin HA0) hoặc hai chuỗi polypeptit (tức là tương ứng với vùng gốc của polypeptit hemagglutinin HA1 trong mối liên hợp với polypeptit hemagglutinin HA2). Theo sáng chế, polypeptit vùng gốc chứa một hoặc nhiều đột biến khi so với phân tử HA kiểu đại, cụ thể là một hoặc nhiều gốc axit amin của HA kiểu đại có thể đã được thay thế bằng các axit amin khác, không xảy ra trong tự nhiên

ở vị trí tương ứng trong HA kiểu đại cụ thể. Polypeptit vùng gốc theo sáng chế có thể còn chứa một hoặc nhiều trình tự liên kết, như được mô tả dưới đây.

Thuật ngữ “vector” chỉ phân tử axit nucleic mà trong đó phân tử axit nucleic thứ hai có thể được chèn vào để đưa vật chủ vào trong khi nó sẽ được sao chép và trong một số trường hợp là được biểu hiện. Nói cách khác, vector có khả năng vận chuyển phân tử axit nucleic mà nó được liên kết. Các vector tách dòng cũng như vector biểu hiện được bao hàm trong thuật ngữ “vector”, như được sử dụng ở đây. Các vector bao gồm, nhưng không chỉ giới hạn ở, plasmid, cosmit, nhiễm sắc thể vi khuẩn nhân tạo (bacterial artificial chromosomes -BAC) và nhiễm sắc thể nấm nhân tạo (yeast artificial chromosomes - YAC) và các vector có nguồn gốc từ thực khuẩn thể hoặc các virus thực vật và động vật (bao gồm người). Các vector bao gồm điểm khởi đầu sao chép được nhận diện bởi vật chủ được đề xuất và trong trường hợp các vector biểu hiện, trình tự khởi đầu phiên mã và các vùng điều hòa khác được nhận diện bằng vật chủ. Một số vector có khả năng tự sao chép trong vật chủ mà chúng được đưa vào (ví dụ, các vector có điểm khởi đầu sao chép của vi khuẩn có thể sao chép trong vi khuẩn). Các vector khác có thể được tích hợp vào bộ gen của vật chủ khi được đưa vào vật chủ, và do đó được sao chép cùng với bộ gen của vật chủ.

Như được sử dụng ở đây, thuật ngữ "kiểu đại" để chỉ virus cúm lưu hành phổ biến trong tự nhiên và gây ra sự bùng phát bệnh dịch điển hình.

Virus cúm có tác động đáng kể đến sức khỏe cộng đồng toàn cầu, gây ra hàng triệu ca bệnh trầm trọng mỗi năm, hàng nghìn ca tử vong, và thiệt hại đáng kể về kinh tế. Vaccine cúm tam liên hiện nay kích thích các kháng thể trung hòa tiềm năng đáp ứng với các chủng trong vaccine và các chủng phân lập có quan hệ gần, nhưng ít khi mở rộng đối với các chủng khác nhau thuộc một typ phụ hoặc các typ phụ khác. Ngoài ra, việc chọn lọc chủng làm vaccine thích hợp gặp phải nhiều thách thức và thường dẫn đến sự phòng ngừa chưa được tối ưu. Hơn nữa, việc dự đoán typ phụ của virus có thể gây đại dịch trong mùa tới, bao gồm thời gian và địa điểm chúng sẽ xuất hiện hiện nay là điều không thể.

Hemagglutinin (HA) là glycoprotein chính của vỏ ngoài từ virus cúm A là đích chủ yếu của các kháng thể trung hòa. Hemagglutinin có hai chức năng chính trong quá trình xâm nhập của virus. Thứ nhất, hemagglutinin làm trung gian cho sự bám dính của virus lên bề mặt của tế bào đích thông qua tương tác với thụ thể axit sialic. Thứ hai, sau quá trình nhập bào của virus, tiếp theo hemagglutinin khởi động quá trình dung hợp của màng virus và màng endosom để giải phóng bộ gen của nó vào tế bào chất của tế bào đích. HA bao gồm vùng ngoài lớn gồm khoảng 500 axit amin được phân cắt bởi

enzym của vật chủ tạo ra 2 polypeptit mà vẫn được liên kết bằng liên kết disulfit. Phần lớn của đoạn đầu tận cùng N (HA1, 320-330 axit amin) tạo ra vùng đầu hình cầu xa tâm màng chứa vị trí liên kết thụ thể và hầu hết các yếu tố quyết định được nhận diện bởi các kháng thể trung hòa virut. Phần đầu tận cùng C nhỏ hơn (HA2, khoảng 180 axit amin) tạo ra cấu trúc tương tự cuống giữ vùng đầu hình cầu gắn vào màng tế bào hoặc màng virut. Mức độ tương đồng về trình tự giữa các typ phụ nhỏ hơn trong các polypeptit HA1 (độ tương đồng 34% - 59% giữa các typ phụ) so với trong polypeptit HA2 (độ tương đồng 51%- 80%). Vùng bảo thủ nhất là trình tự xung quanh vị trí phân cắt, cụ thể là 23 axit amin đầu tận cùng N của HA2, được bảo thủ trong số các typ phụ của virut cúm A (Lorieau et al., 2010). Một phần của vùng này được bọc lộ dưới dạng vòng trên bề mặt trong phân tử HA tiền thân (HA0), nhưng trở nên không thể xâm nhập được khi HA0 được cắt thành HA1 và HA2.

Hầu hết các kháng thể trung hòa liên kết với các vòng quanh vị trí liên kết thụ thể và cản trở sự liên kết với thụ thể và bám dính. Do các vòng này có tính biến đổi cao, hầu hết kháng thể hướng đích các vùng này là đặc hiệu chủng, điều này giải thích tại sao vaccin hiện nay kích thích hệ miễn dịch giới hạn, đặc hiệu chủng này. Tuy nhiên, gần đây, kháng thể đơn dòng hoàn toàn của người hemagglutinin của kháng virut cúm có khả năng trung hòa chéo ở phạm vi rộng được tạo ra. Phân tích cấu trúc và chức năng đã cho thấy rằng các kháng thể này cản trở quá trình dung hợp của màng và nhằm kháng epitop bảo thủ cao trong vùng gốc của protein HA của virut cúm (Throsby et al., 2008; Ekiert et al. 2009, WO 2008/028946, WO 2010/130636).

Theo sáng chế polypeptit vùng gốc của HA mới đã được thiết kế chứa các epitop này để tạo ra vaccin trên cơ sở epitop phổ quát gây cảm ứng sự bảo vệ kháng lại phổ rộng các chủng virut cúm. Về bản chất, phần quyết định miễn dịch và có khả năng biến đổi cao, tức là vùng đầu, trước tiên được loại từ phân tử HA chiều dài đủ để tạo ra polypeptit vùng gốc, còn được gọi là mini-HA. Trong phương pháp này đáp ứng miễn dịch sẽ được chuyển hướng đến vùng gốc trong đó có các epitop cho các kháng thể trung hòa phổ rộng. Các kháng thể trung hòa phổ rộng được đề cập trên đây được sử dụng để thăm dò sự cuộn gấp chính xác của các phân tử mới được tạo ra, và để xác nhận sự có mặt của các epitop trung hòa.

Polypeptit vùng gốc theo sáng chế có khả năng trình diện các epitop bảo thủ của phân tử HA vùng cuống gần gốc màng cho hệ miễn dịch khi không có mặt các epitop trội mà có trong vùng đầu ở xa màng. Cuối cùng, một phần của trình tự ban đầu của protein HA0 tạo ra vùng đầu được loại bỏ và được nối lại, hoặc theo cách trực tiếp, hoặc theo một số phương án, bằng cách đưa vào trình tự nối ngắn linh động ('cầu

nổi') để khôi phục tính liên tục của chuỗi polypeptit. Trình tự polypeptit thu được còn được cải biến bằng cách đưa vào các đột biến đặc hiệu làm ổn định cấu trúc ba chiều tự nhiên của phần còn lại của phân tử HA0.

Sáng chế do đó đề xuất các polypeptit chứa (a) vùng HA1 của hemagglutinin của virus cúm mà bao gồm đoạn gốc đầu tận cùng N của HA1, được liên kết cộng hóa trị bằng trình tự liên kết gồm 0-50 gốc axit amin với đoạn gốc đầu tận cùng C của HA1, và (b) vùng HA2 của hemagglutinin của virus cúm, trong đó trên hoặc nhiều axit amin trong vùng HA2 được làm đột biến. Trong polypeptit theo sáng chế, vùng HA2 do đó chứa một hoặc nhiều đột biến so với vùng HA2 của hemagglutinin của virus cúm kiểu đại mà polypeptit vùng gốc của HA dựa vào.

Các polypeptit vùng gốc hemagglutinin của virus cúm này dựa trên HA của typ phụ của virus cúm A mà thường được sử dụng trong vacxin virus cúm cho người. Theo các phương án ưu tiên, polypeptit vùng gốc dựa trên HA của virus cúm chứa HA thuộc các typ phụ H1, H5 và/hoặc H3.

Cụ thể sáng chế đề xuất các polypeptit vùng gốc hemagglutinin của virus cúm chứa (a) vùng HA1 của hemagglutinin của virus cúm mà bao gồm đoạn gốc đầu tận cùng N của HA1, được liên kết cộng hóa trị bằng trình tự liên kết gồm 0-50 gốc axit amin với đoạn gốc đầu tận cùng C của HA1, và (b) vùng HA2 của hemagglutinin của virus cúm, trong đó polypeptit vùng gốc hemagglutinin kháng lại sự phân cắt bằng proteaza ở vị trí nối giữa HA1 và HA2, và trong đó một hoặc nhiều axit amin trong trình tự axit amin nối chuỗi cuộn xoắn A và chuỗi cuộn xoắn CD của HA2 đã được gây đột biến so với vùng HA2 của virus cúm kiểu đại. Tốt hơn, vùng HA1 và HA2 có nguồn gốc từ typ phụ virus cúm A được chọn từ nhóm bao gồm H1, H5 và H3.

Polypeptit theo sáng chế do đó chứa một hoặc nhiều đột biến trong trình tự axit amin của HA2 nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD, như được chỉ ra trên FIG. 1. Theo một số phương án, một hoặc nhiều axit amin kỵ nước trong trình tự axit amin của HA2 này đã được thay thế bằng các axit amin ưa nước, như các axit amin phân cực và/hoặc tích điện, hoặc axit amin linh động glyxin (G).

Polypeptit theo sáng chế không chứa HA1 chiều dài đủ.

Theo một số phương án, các polypeptit gây miễn dịch hầu như nhỏ hơn HA0, tốt hơn không có toàn bộ hoặc hầu như toàn bộ vùng đầu hình cầu của HA. Tốt hơn, các polypeptit gây miễn dịch có chiều dài không nhiều hơn 360, tốt hơn là không nhiều hơn 350, 340, 330, 320, 310, 305, 300, 295, 290, 285, 280, 275, hoặc 270 axit amin. Theo một số phương án, các polypeptit gây miễn dịch có chiều dài khoảng 250

đến 350, tốt hơn là từ 260 đến 340, tốt hơn là từ 270 đến 330, tốt hơn là từ 270 đến 330 axit amin.

Theo một số phương án, polypeptit còn chứa một hoặc nhiều đột biến khác trong vùng HA1 và/hoặc HA2, khi so với trình tự axit amin của HA mà các vùng HA1 và HA2 thu được từ đó. Do đó, độ ổn định của các polypeptit vùng gốc còn được gia tăng.

Theo sáng chế, “đoạn gốc đầu tận cùng N của HA1” chỉ đoạn polypeptit mà tương ứng với phần đầu tận cùng amin của vùng HA1 of của phân tử hemagglutinin (HA) của virus cúm. Theo một số phương án, đoạn polypeptit đầu tận cùng N của HA1 chứa axit amin từ vị trí 1 đến vị trí x của vùng HA1, trong đó axit amin ở vị trí x là gốc axit amin nằm trong HA1. Thuật ngữ “đoạn đầu tận cùng C của HA1” chỉ đoạn polypeptit mà tương ứng với phần đầu tận cùng carboxy của vùng HA1 của hemagglutinin của virus cúm. Theo một số phương án, đoạn polypeptit đầu tận cùng C của HA1 chứa axit amin từ vị trí y đến và bao gồm axit amin đầu tận cùng C của vùng HA1, trong đó axit amin ở vị trí y là gốc axit amin nằm trong HA1. Theo sáng chế y lớn hơn x, do đó một đoạn của vùng HA1 giữa đoạn đầu tận cùng N của HA1 và đoạn đầu tận cùng C của HA1, tức là giữa axit amin ở vị trí x và axit amin ở vị trí y của HA1, được loại bỏ, và theo một số phương án, được thay thế bằng trình tự nối.

Theo một số phương án đoạn gốc đầu tận cùng N của HA1 gồm axit amin 1-x của HA1, và đoạn gốc đầu tận cùng C của HA1 gồm các axit amin y-điểm cuối của HA1. Do đó, theo một số phương án, việc loại bỏ đoạn HA1 chứa trình tự axit amin từ axit amin ở vị trí x+1 đến và bao gồm axit amin ở vị trí y-1.

Theo một số phương án, polypeptit không chứa trình tự tín hiệu. Do đó theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm axit amin p-x của HA1, trong đó p là axit amin đầu tiên của phân tử HA trưởng thành (ví dụ p=18 trong trường hợp trình tự SEQ ID NO: 1). Người có hiểu biết trung bình trong lĩnh vực sẽ có thể tạo ra polypeptit được mô tả ở đây không có các peptit tín hiệu (ví dụ axit amin 1-17 của trình tự SEQ ID NO: 1). Theo một số phương án, polypeptit theo sáng chế chứa trình tự nội bào của HA và vùng xuyên màng. Theo các phương án khác, polypeptit theo sáng chế không chứa trình tự nội bào của HA và vùng xuyên màng. Theo một số phương án, trình tự nội bào và xuyên màng, ví dụ trình tự axit amin từ vị trí (hoặc vị trí tương đương của) 523, 524, 525, 526, 527, 526, 528, 529, hoặc 530 của vùng HA2 đến đầu tận cùng C của vùng HA2 được loại bỏ.

Theo sáng chế, polypeptit vùng gốc của hemagglutinin này kháng với sự phân cắt bằng proteaza ở vị trí nối giữa HA1 và HA2. Người có hiểu biết trung bình trong

lĩnh vực biết được rằng trình tự Arg (R) - Gly (G) nối HA1 và HA2 là vị trí nhận biến của các proteaza tương tự trypsin và trypsin và thường được cắt để hoạt hóa hemagglutinin. Do polypeptit vùng gốc của HA được mô tả ở đây không được hoạt hóa, các polypeptit vùng gốc hemagglutinin của virus cúm theo sáng chế kháng với sự phân cắt bằng proteaza. Theo sáng chế, do đó vị trí phân cắt bằng proteaza được loại bỏ hoặc vị trí của proteaza nối HA1 và HA2 được gây đột biến thành trình tự mà kháng với sự phân cắt bằng proteaza.

Theo một số phương án, gốc axit amin đầu tận cùng C của đoạn gốc đầu tận cùng C của HA1 là axit amin bất kỳ không phải arginin (R) hoặc lysin (K). Theo một số phương án, axit amin đầu tận cùng C của HA1 là glutamin (Q), serin (S), treonin (T), asparagin (N), axit aspartic (D) hoặc axit glutamic (E). Theo một số phương án, gốc axit amin đầu tận cùng C của đoạn gốc đầu tận cùng C của HA1 là glutamin (Q).

Theo một số phương án, polypeptit được glycosyl hóa.

Các polypeptit vùng gốc hemagglutinin của virus cúm này có thể dựa trên HA của hemagglutinin của virus cúm A bất kỳ có trong tự nhiên thuộc typ phụ được sử dụng trong vaccin cúm cho người. Typ phụ của virus cúm A thường được sử dụng trong vaccin cúm là virus cúm A thuộc các typ phụ H1, H3 hoặc H5. Với “dựa trên” có nghĩa là các đoạn đầu tận cùng N, và/hoặc các đoạn đầu tận cùng C của vùng HA1 và/hoặc vùng HA2 có độ tương đồng về trình tự axit amin ít nhất là 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, hoặc 99% với các đoạn đầu tận cùng N và/hoặc các đoạn đầu tận cùng C tương ứng của HA1 và/hoặc vùng HA2 của hemagglutinin của virus cúm bất kỳ có trong tự nhiên thuộc typ phụ H1, H3 và/hoặc H5 mà người có hiểu biết trung bình trong lĩnh vực đã biết hoặc được phát hiện sau. Theo một số phương án, các polypeptit vùng gốc hemagglutinin của virus cúm này dựa trên hemagglutinin của virus cúm thuộc virus cúm A nhóm 1. Theo một số phương án, các polypeptit vùng gốc hemagglutinin của virus cúm này dựa trên hemagglutinin của virus cúm thuộc virus cúm A nhóm 2. Theo một số phương án, polypeptit vùng gốc hemagglutinin của virus cúm là polypeptit lai hoặc thể khảm mà bao gồm hoặc gồm các đoạn và/hoặc các vùng từ nhiều chủng hoặc typ phụ của virus cúm khác nhau. Ví dụ, polypeptit vùng gốc hemagglutinin của virus cúm có thể chứa các đoạn gốc đầu tận cùng N của HA1 và đoạn gốc đầu tận cùng C của HA1 và/hoặc các vùng của HA2 từ các typ phụ HA khác nhau của virus cúm A.

Theo một số phương án, polypeptit dựa trên HA của H1. Theo phương án cụ thể, polypeptit chứa các vùng gốc hemagglutinin từ hoặc dựa trên HA của virus cúm A chứa HA của typ phụ H1, như từ virus cúm A/Brisbane/59/2007 (H1N1) (SEQ ID

NO:1), như được mô tả dưới đây. Người có hiểu biết trung bình trong lĩnh vực cần phải hiểu rằng virus cúm A chứa HA của typ phụ H1 khác cũng có thể được sử dụng theo sáng chế. Theo một số phương án, polypeptit chứa các vùng gốc hemagglutinin dựa trên HA của virus cúm A H1 được chọn từ Bảng 7.

Theo một số phương án, polypeptit chứa đoạn polypeptit đầu tận cùng N của HA1 chứa axit amin từ vị trí 1 đến vị trí x của HA của vùng H1, trong đó x là axit amin bất kỳ nằm giữa axit amin ở vị trí 46 và axit amin ở vị trí 60, như axit amin ở vị trí 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, hoặc 59, tốt hơn trong đó x là 52, 53, 55 hoặc 59. Tốt hơn, polypeptit chứa đoạn đầu tận cùng N của HA1 không có trình tự tín hiệu, tức là đoạn đầu tận cùng N của HA1 chứa axit amin từ vị trí 18 (ví dụ đối với HA của H1, như SEQ ID NO: 1), hoặc vị trí tương đương trong các chủng virus cúm H1 khác, đến vị trí x của vùng HA1. Theo một số phương án, đoạn đầu tận cùng N của HA1 do đó chứa axit amin từ vị trí p (trong đó p=18 đối với HA của H1 trong trình tự SEQ ID NO: 1 hoặc vị trí tương đương trên HA của các H1 khác), đến vị trí x của vùng HA1.

Theo một số phương án, đoạn polypeptit đầu tận cùng C của HA1 chứa axit amin từ vị trí y đến và bao gồm axit amin đầu tận cùng C của HA của vùng H1, trong đó y là axit amin bất kỳ nằm giữa axit amin ở vị trí 290 và axit amin ở vị trí 325 của HA của H1, tốt hơn trong đó y là 291, 303, 318, hoặc 321. Theo sáng chế, vùng HA2 chứa một hoặc nhiều đột biến trong trình tự axit amin của HA2 nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD (FIG. 1). Theo một số phương án, một hoặc nhiều axit amin kỵ nước trong trình tự axit amin của HA2 này đã được thay thế bằng các axit amin ưa nước, như axit amin phân cực và/hoặc tích điện. Theo một số phương án (ví dụ đối với HA của H1, như SEQ ID NO: 1), trình tự axit amin của HA2 nối gốc ở đầu tận cùng C của chuỗi xoắn A gốc đầu tận cùng N của chuỗi cuộn xoắn CD chứa trình tự axit amin giữa các gốc 402-418 của HA2 của virus cúm. Theo một số phương án, trình tự axit amin của HA2 nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc đầu tận cùng N của chuỗi cuộn xoắn CD chứa trình tự axit amin MNTQFTAVGKEFN(H/K)LE(K/R) (SEQ ID NO: 17).

Theo một số phương án, x là 59 và y là 291.

Theo một số phương án, x là 52 và y là 321.

Theo một số phương án x là 53 và y là 303.

Theo một số phương án x là 55 và y là 318.

Theo một phương án, trình tự axit amin nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD tương ứng với trình tự axit amin

giữa axit amin ở vị trí 402 và axit amin ở vị trí 418 của HA2 của trình tự SEQ ID NO: 1, trong đó polypeptit chứa một hoặc nhiều đột biến trong trình tự axit amin kéo dài từ axit amin 402 đến 418 của trình tự SEQ ID NO: 1. Trình tự axit amin giữa gốc 402-418 của HA của virus cúm thuộc typ huyết thanh H1 chứa trình tự axit amin MNTQFTAVGKEFN(H/K)LE(K/R) (SEQ ID NO: 17). Theo một số phương án, trình tự axit amin giữa gốc 402-418 của HA của virus cúm thuộc typ huyết thanh H1 chứa trình tự axit amin MNTQX₁TAX₂GKEX₃N(H/K)X₄E(K/R).

Theo một số phương án polypeptit do đó chứa một hoặc nhiều đột biến trong vùng HA2 của H1 như được chỉ ra trên Bảng 6. Theo một số phương án, một hoặc nhiều axit amin ở vị trí 406, 409, 413 và 416, tức là một hoặc nhiều axit amin trong số các axit amin X₁, X₂, X₃ và X₄ được làm đột biến (đánh số đề cập đến trình tự SEQ ID NO: 1). Theo một số phương án, axit amin ở vị trí 406, tức là X₁ được làm biến đổi thành axit amin được chọn từ nhóm bao gồm S, T, N, Q, R, H, K, D, E, và G, tốt hơn là S. Theo một số phương án, axit amin ở vị trí 409, tức là X₂ được làm biến đổi thành axit amin được chọn từ nhóm bao gồm S, T, N, Q, R, H, K, D, E, và G, tốt hơn là T, Q hoặc G. Theo một số phương án, axit amin ở vị trí 413, tức là X₃ được làm biến đổi thành axit amin được chọn từ nhóm bao gồm S, T, N, Q, R, H, K, D, E, G, tốt hơn là S. Theo một số phương án, axit amin ở vị trí 416, tức là X₄ được làm biến đổi thành axit amin được chọn từ nhóm bao gồm S, T, N, Q, R, H, K, D, E, G, tốt hơn là S. Cũng có thể là sự kết hợp của các đột biến này.

Theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm các gốc axit amin 1-59 của HA1, và đoạn gốc đầu tận cùng C của HA1 chứa các gốc axit amin 291-343 của HA1 trong đó axit amin ở vị trí 343, tức là R343, được làm đột biến và là axit amin không phải R, tốt hơn là glutamin (Q). Theo một số phương án, đoạn đầu tận cùng N của HA1 gồm các gốc axit amin 1-59 của HA1 và đoạn đầu tận cùng C của HA1 gồm các gốc axit amin 291-343 của HA1. Lưu ý rằng số lượng axit amin dựa trên số lượng axit amin trong HA của H10, cụ thể là số lượng axit amin của chủng virus cúm H1N1 A/Brisbane/59/2007 (SEQ ID NO: 1). Lưu ý rằng do các trình tự HA của các typ phụ/chủng virus cúm khác nhau có thể có các đoạn chèn vào hoặc đoạn mất đi trong vùng đầu so với nhau, nên việc đánh số không phải luôn luôn là như nhau. Người có hiểu biết trung bình trong lĩnh vực có thể xác định các vị trí axit amin tương đương trong các trình tự HA của các chủng virus cúm và/hoặc các typ phụ khác bằng cách sắp xếp và so sánh trình tự.

Theo một số phương án, đoạn polypeptit đầu tận cùng N của HA1 không chứa trình tự tín hiệu. Theo các phương án ưu tiên đoạn gốc đầu tận cùng N của HA1 gồm

axit amin từ vị trí 18 đến vị trí 59 của vùng HA1. Theo một số phương án, đoạn đầu tận cùng N của HA1 gồm axit amin 18-59 của vùng HA1.

Theo một số phương án, polypeptit theo sáng chế chứa một hoặc nhiều đột biến khác, tức là các đột biến thay thế axit amin trong vùng HA1 và/hoặc vùng HA2. Theo một số phương án, vùng HA1 do đó còn chứa một hoặc nhiều đột biến trong số các đột biến sau: L58T, V314T và I316T. Một lần nữa cần lưu ý rằng số lượng axit amin dựa trên số lượng axit amin trong HA của H10, cụ thể là số lượng axit amin của chủng virus cúm H1N1 A/Brisbane/59/2007 (SEQ ID NO: 1). Người có hiểu biết trung bình trong lĩnh vực có thể xác định axit amin tương đương trong HA của virus cúm H1 và do đó có thể xác định đột biến tương đương.

Theo phương án cụ thể, vùng HA1 chứa đột biến L58T, V314T, và I316T, và vùng HA2 chứa một hoặc nhiều đột biến trong số các đột biến sau: F406S, V409T, và L416S.

Theo một số phương án, vùng HA1 còn chứa đột biến K321C và/hoặc vùng HA2 còn chứa một hoặc nhiều đột biến trong số các đột biến sau: Q405C, F413C, E421C, và Y502S.

Theo phương án cụ thể, vùng HA1 chứa đột biến L58T, V314T, I316T, và K321C và vùng HA2 chứa đột biến: Q405C, F406S, V409T, và L416S.

Theo phương án cụ thể, vùng HA1 chứa đột biến L58T, V314T, và I316T, và vùng HA2 chứa đột biến: F406S, V409T, F413C, L416S và E421C.

Theo phương án cụ thể, vùng HA1 chứa đột biến L58T, V314T, và I316T, và vùng HA2 chứa đột biến: F406S, V409T, L416S, và Y502S.

Theo phương án cụ thể, vùng HA1 chứa đột biến L58T, V314T, I316T, và K321C và vùng HA2 chứa đột biến: Q405C, F406S, V409T, F413C, L416S và E421C.

Theo phương án cụ thể, vùng HA1 chứa đột biến L58T, V314T, I316T, và K321C và vùng HA2 chứa đột biến: Q405C, F406S, V409T, F413C, L416S, E421C và Y502S.

Theo các phương án khác, vùng HA2 còn chứa một hoặc nhiều đột biến M420I và V421I, hoặc đột biến tương đương.

Theo phương án cụ thể, vùng HA1 chứa đột biến L58T, V314T, và I316T, và vùng HA2 chứa một hoặc nhiều đột biến trong số các đột biến sau: F406S, V409T, L416S, M420I và V421I.

Theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm các gốc axit amin 1-52 của HA1, tốt hơn là các gốc axit amin 18-52 của HA1, và đoạn gốc đầu tận

cùng C của HA1 chứa các gốc axit amin 321-343 của HA1, trong đó axit amin ở vị trí 343, tức là R343, được làm đột biến và là axit amin không phải R, tốt hơn là glutamin (Q), trong đó vùng HA2 chứa đột biến F406S, V409T, L416S, M420I và V421I. Theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm các gốc axit amin 1-52 của HA1, tốt hơn các gốc axit amin 18-52 của HA1, và đoạn gốc đầu tận cùng C của HA1 gồm các gốc axit amin 321-343 của HA1.

Theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm các gốc axit amin 1-53 của HA1, tốt hơn là các gốc axit amin 18-53 của HA1, và đoạn gốc đầu tận cùng C của HA1 chứa các gốc axit amin 303-343 của HA1, trong đó axit amin ở vị trí 343, tức là R343, được làm đột biến và là axit amin không phải R, tốt hơn là glutamin (Q). Theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm các gốc axit amin 1-53 của HA1, tốt hơn là các gốc axit amin 18-53 của HA1, và đoạn gốc đầu tận cùng C của HA1 gồm các gốc axit amin 303-343 của HA1. Theo phương án cụ thể, vùng HA1 chứa đột biến V314T và I316T, và vùng HA2 chứa một hoặc nhiều đột biến trong số các đột biến sau: F406S, V409T, L416S, M420I và V421I. Theo phương án được ưu tiên, polypeptit chứa trình tự axit amin nêu trong SEQ ID NO: 11.

Theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm các gốc axit amin 1-55 của HA1, tốt hơn là các gốc axit amin 18-55 của HA1, và đoạn gốc đầu tận cùng C của HA1 chứa các gốc axit amin 318-343 của HA1, trong đó axit amin ở vị trí 343, tức là R343, được làm đột biến và là axit amin không phải R, tốt hơn là glutamin (Q). Theo một số phương án, đoạn gốc đầu tận cùng N của HA1 gồm các gốc axit amin 1-55 của HA1, tốt hơn là các gốc axit amin 18-55 của HA1, và đoạn gốc đầu tận cùng C của HA1 gồm các gốc axit amin 318-343 của HA1. Theo một phương án, vùng HA2 chứa đột biến F406S, V409T, L416S, M420I và V421I.

Theo một số phương án, polypeptit còn chứa đột biến R324C trong vùng HA1 và T436C trong vùng HA2.

Theo phương án cụ thể, vùng HA1 chứa đột biến L58T, V314T, I316T, và R324C và vùng HA2 chứa một hoặc nhiều đột biến trong số các đột biến sau: F406S, V409T, L416S, M420I, V421I và T436C.

Theo một phương án, vùng HA1 chứa đột biến R324C, và vùng HA2 chứa đột biến F406S, V409T, L416S, M420I, V421I và T436C.

Theo phương án khác, vùng HA1 chứa đột biến V314T, I316T và R324C, và vùng HA2 chứa một hoặc nhiều đột biến trong số các đột biến sau: F406S, V409T, L416S, M420I, V421I và T436C.

Theo một phương án, vùng HA1 chứa đột biến R324C, và vùng HA2 chứa đột biến F406S, V409T, L416S, M420I, V421I và T436C.

Theo một số phương án, polypeptit theo sáng chế chứa trình tự nội bào của HA và vùng xuyên màng. Theo các phương án khác, các trình tự nội bào và xuyên màng, ví dụ trình tự axit amin từ vị trí (hoặc vị trí tương đương của) 523, 524, 525, 526, 527, 526, 528, 529, hoặc 530 của vùng HA2 đến đầu tận cùng C của vùng HA2 (đánh số theo trình tự SEQ ID NO: 1) được loại bỏ. Theo một số phương án, polypeptit được làm ổn định hơn nữa bằng cách đưa vào trình tự đã biết để tạo cấu trúc trime, tức là AYVRKDGGEWVLL (SEQ ID NO:143) (trình tự ‘cuộn gấp’), tùy ý được nối thông qua cầu nối. Cầu nối này có thể tùy ý chứa vị trí phân cắt để xử lý sau đó theo quy trình mà người có hiểu biết trung bình trong lĩnh vực đã biết. Để dễ dàng tinh chế dạng hòa tan có thể bổ sung thêm trình tự đuôi, ví dụ đuôi his (HHHHHHH) được nối thông qua cầu nối ngắn, ví dụ EGR. Theo một số phương án cầu nối và trình tự đuôi his được thêm vào không có trình tự cuộn gấp được thể hiện.

Theo một số phương án, trình tự axit amin từ vị trí (hoặc vị trí tương đương của) 530 của vùng HA2 đến đầu tận cùng C của vùng HA2 (đánh số theo trình tự SEQ ID NO: 1) được loại bỏ. Theo một số phương án, trình tự nội bào và xuyên màng được thay thế bằng trình tự axit amin AGRHHHHHHH (SEQ ID NO: 81) hoặc SGRSLVPRGSPGSGYIPEAPRDGQAYVRKDGGEWVLLSTFLGHHHHHHH (SEQ ID NO: 82).

Theo một số phương án, polypeptit liên kết có chọn lọc với kháng thể CR6261 và/hoặc CR9114. Theo một phương án, polypeptit không liên kết với kháng thể CR8057. Theo một phương án, CR6261 chứa vùng biến đổi của chuỗi nặng chứa trình tự axit amin nêu trong SEQ ID NO: 20 và vùng biến đổi của chuỗi nhẹ chứa trình tự axit amin nêu trong SEQ ID NO: 21; CR9114 chứa vùng biến đổi của chuỗi nặng chứa trình tự axit amin nêu trong SEQ ID NO: 18 và vùng biến đổi của chuỗi nhẹ chứa trình tự axit amin nêu trong SEQ ID NO: 19. Theo một phương án, CR8057 chứa vùng biến đổi của chuỗi nặng chứa trình tự axit amin nêu trong SEQ ID NO: 22 và vùng biến đổi của chuỗi nhẹ chứa trình tự axit amin nêu trong SEQ ID NO: 23.

Như được mô tả trên đây, polypeptit chứa vùng HA1 của hemagglutinin của virus cúm mà bao gồm đoạn gốc đầu tận cùng N của HA1 được liên kết cộng hóa trị bằng trình tự nối gồm 0- 50 các gốc axit amin với đoạn gốc đầu tận cùng C của HA1. Trình tự nối không có trong tự nhiên, hoặc HA kiểu dại. Theo một số phương án, cầu nối là peptit mà bao gồm một gốc axit amin, hai gốc axit amin hoặc ít hơn, ba gốc axit amin hoặc ít hơn, bốn gốc axit amin hoặc ít hơn, năm gốc axit amin hoặc ít hơn, mười

gốc axit amin hoặc ít hơn, 15 gốc axit amin hoặc ít hơn, hoặc 20 gốc axit amin hoặc ít hơn hoặc 30 gốc axit amin hoặc ít hơn hoặc 40 gốc axit amin hoặc ít hơn hoặc 50 gốc axit amin hoặc ít hơn. Theo phương án cụ thể, trình tự nối là trình tự được chọn từ nhóm bao gồm G, GS, GGG, GSG, GSA, GSGS, GSAG, GGGG, GSAGS, GSGSG, GSAGSA, GSAGSAG, và GSGSGSG.

Sáng chế còn đề cập đến các phương pháp tạo ra polypeptit theo sáng chế, cụ thể là để tạo ra polypeptit vùng gốc HA của H1 theo sáng chế, cũng như polypeptit có thể thu được hoặc thu được bằng các phương pháp này. Theo một số phương án, các phương pháp này bao gồm các bước:

(a) Cung cấp trình tự axit amin HA0 của virus cúm, cụ thể là trình tự axit amin HA0 của virus cúm thuộc typ huyết thanh H1;

(b) Loại bỏ vị trí cắt giữa HA1 và HA2, tốt hơn là bằng cách gây đột biến axit amin đầu tận cùng C của HA1 thành axit amin khác với arginin (R) hoặc lysin (K);

(c) Loại bỏ trình tự axit amin của vùng đầu hình cầu từ trình tự HA0; Bước này được thực hiện bằng cách loại bỏ một đoạn của vùng HA1 giữa axit amin ở vị trí x và axit amin ở vị trí y, và nối lại đoạn đầu tận cùng N (kéo dài từ axit amin ở vị trí 1 đến và bao gồm axit amin ở vị trí x của HA1) và đoạn đầu tận cùng C của HA1 (kéo dài từ axit amin y đến axit amin đầu tận cùng C của HA1), do đó thu được, tùy ý thông qua trình tự nối gồm 0-50 axit amin. Theo một số phương án, x là axit amin ở vị trí bất kỳ giữa các vị trí 46 và 60, tốt hơn axit amin ở vị trí 52, 53, 55 hoặc 59 của HA1 và trong đó y là axit amin ở vị trí bất kỳ giữa các vị trí 290 và 325, tốt hơn axit amin ở vị trí 291, 303, 318, hoặc 321 của HA1. Một lần nữa, việc đánh số được sử dụng đề cập đến trình tự SEQ ID NO:1. Người có hiểu biết trung bình trong lĩnh vực cần phải hiểu rằng trình tự dẫn đầu (hoặc trình tự tín hiệu) mà điều khiển việc vận chuyển của protein trong quá trình sản sinh (ví dụ tương ứng với axit amin 1-17 của trình tự SEQ ID NO: 1), thường sẽ không có mặt trong polypeptit cuối cùng, ví dụ được sử dụng trong vaccine. Theo một số phương án, polypeptit theo sáng chế do đó chứa đoạn đầu tận cùng N của HA1 không có trình tự dẫn đầu.

(d) Làm tăng độ ổn định của cấu hình trước khi dung hợp và loại bỏ việc làm ổn định của cấu hình sau khi dung hợp của HA cải biến, tốt hơn là bằng cách đưa một hoặc nhiều đột biến vào trong trình tự axit amin nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD, tốt hơn là trong trình tự axit amin kéo dài từ axit amin 402-418 của trình tự SEQ ID NO: 1 cụ thể là chứa trình tự axit amin MNTQFTAVGKEFN(H/K)LE(K/R) (SEQ ID NO: 17). Đột biến tốt hơn là đột biến thay thế các gốc axit amin kỵ nước thành các gốc axit amin ưa nước.

(e) Đưa một hoặc nhiều cầu disulfit vào polypeptit vùng gốc của HA.

Theo sáng chế, việc loại bỏ vị trí cắt giữa HA1 và HA2 có thể đạt được bằng đột biến R (một số ít trường hợp là K) thành Q ở vị trí P1 (xem ví dụ Sun et al, 2010 để biết giải thích về danh pháp của vị trí cắt (vị trí 343 trong trình tự SEQ ID NO: 1). Đột biến thành Q được ưu tiên nhưng S, T, N, D hoặc E là các dạng thay thế.

Việc loại bỏ vùng đầu có thể đạt được, ví dụ bằng cách loại bỏ axit amin 53 đến 320 từ SEQ ID NO: 1, hoặc ở các vị trí tương đương trong HA từ virus cúm khác. Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp ví dụ như Clustal hoặc Muscle. Các phần còn lại của trình tự có thể được nối trực tiếp, hoặc theo cách khác đưa vào cầu nối linh động. Trình tự cầu nối có thể có chiều dài từ 1 đến 50 axit amin. Được ưu tiên là các cầu nối linh động có chiều dài giới hạn (nhỏ hơn hoặc bằng 10 axit amin), ví dụ GGG, GGGG, GSA, GSAG, GSAGSA, GSAGSAG hoặc tương tự. Chiều dài của đoạn bị loại bỏ cũng có thể thay đổi, ví dụ bằng cách bắt đầu loại bỏ ở (tương đương của) vị trí (x), ví dụ ở vị trí 54, 55, 56, 57 hoặc 58, hoặc làm tăng chiều dài của đoạn bị loại bỏ, bằng cách cắt ở vị trí 47, 48, 49, 50, 51, hoặc 52. Tương tự, axit amin cuối cùng bị loại bỏ có thể ở (tương đương của) vị trí (y), như 315, 316, 317, 318 hoặc 319, hoặc làm tăng chiều dài của đoạn bị loại bỏ ở (tương đương của) vị trí 321, 322, 323, 324, hoặc 325. Quan trọng là nhận diện được sự thay đổi về chiều dài của đoạn được loại bỏ có thể ở trong phần được bù bằng cách làm cho phù hợp chiều dài của trình tự cầu nối, tức là đoạn loại bỏ lớn hơn có thể được làm phù hợp với cầu nối dài hơn và ngược lại. Các polypeptit này cũng có thể được bao gồm trong sáng chế.

Theo sáng chế, độ hòa tan của vòng giữa chuỗi cuộn xoắn A và chuỗi cuộn xoắn CD được gia tăng. Vòng này được tạo thành bởi (tương đương của) các gốc 402 đến 418 trong chủng H1 A/Brisbane/59/2007 (SEQ ID NO: 1). Do đó, độ ổn định của cấu hình trước khi dung hợp được tăng lên và cấu hình sau khi dung hợp của HA được cải biến được làm mất ổn định. Vòng này có tính bảo thủ cao trong trình tự H1, như được thể hiện trong bảng 6 dưới đây. Điều này có thể đạt được chẳng hạn bằng cách thay thế axit amin I, L, F hoặc V trong vòng này bằng các bản sao ưa nước. Vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp ví dụ như Clustal hoặc Muscle. Đột biến thành glyxin làm mất ổn định cấu hình sau khi dung hợp do độ linh động cao của axit amin này dẫn đến sự giảm độ ổn định của chuỗi cuộn xoắn sau dung hợp được tạo thành bởi phần này của trình tự HA. Trình tự liên ứng mô

tả vòng giữa gốc 402-418 của HA của virus cúm thuộc typ huyết thanh H1 là (SEQ ID NO: 17) MNTQFTAVGKEFN(H/K)LE(K/R) . Trong polypeptit theo sáng chế axit amin ở các vị trí 406, 409, 413 và/hoặc 416 (hoặc vị trí tương đương của chúng, như được xác định từ sắp xếp và so sánh trình tự) là axit amin phân cực (S,T,N,Q), tích điện (R,H,K,D,E) hoặc linh động (G). Cũng có thể là tổ hợp của các đột biến này ở các vị trí này, ví dụ F406S, V409T, L416S. Trong một số trường hợp đột biến để phục hồi axit amin liên ứng được ưu tiên, ví dụ trong đó V hoặc M ở vị trí 404 (thành T), V ở 408 (thành A) hoặc 410 (thành G) hoặc I ở 414 (thành N); tỷ lệ đột biến trong trình tự ở các axit amin này là rất thấp. Tổng quan về đột biến được mô tả trên đây mà xác định đặc tính của polypeptit theo sáng chế được đưa ra trong bảng 6.

Theo sáng chế, một hoặc nhiều cầu disulfit được đưa vào trong polypeptit vùng gốc, tốt hơn giữa axit amin ở (hoặc vị trí tương đương của) vị trí 324 và 436 trong chủng H1 A/Brisbane/59/2007. Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp như Clustal, Muscle etc. Cầu disulfit được thiết kế được tạo ra bằng cách gây đột biến ít nhất một (nếu gốc còn lại là cystein rồi), nhưng thường là hai gốc gần nhau về không gian với cystein, mà sẽ tạo thành liên kết cộng hóa trị một cách tự phát hoặc bằng phản ứng oxi hóa hoạt hóa giữa các nguyên tử lưu huỳnh của các gốc này.

HA tự nhiên tồn tại ở dạng trime trên bề mặt tế bào. Hầu hết các tương tác giữa từng monome mà giữ dạng trime cùng nằm trong vùng đầu. Sau khi loại bỏ vùng đầu, do đo cấu trúc bậc ba được làm mất ổn định và vì vậy củng cố các tương tác giữa các monome trong phân tử bị cắt đứt sẽ làm tăng độ ổn định. Trong vùng gốc, trime hóa được trung gian bởi sự hình thành của môtip cuộn xoắn dạng trime. Bằng cách củng cố môtip này có thể tạo ra dạng trime thích hợp hơn. Theo sáng chế, trình tự liên ứng để tạo ra dạng trime xoắn cuộn, ví dụ IEAIEKKIEAIEKKIE (SEQ ID NO: 83), có thể được đưa vào trong polypeptit theo sáng chế ở (tương đương của) vị trí 418 đến 433. Theo một số phương án, trình tự MKQIEDKIEEIESKQ (SEQ ID NO: 84), có nguồn gốc từ GCN4 và còn được biết là để trime hóa được đưa vào ở (tương đương của) vị trí 419-433. Theo một số phương án, mặt trung gian của trime được làm ổn định bằng cách cải biến M420, L423, V427, G430 thành isoleuxin.

Theo một số phương án, polypeptit theo sáng chế chứa trình tự nội bào của HA của H1 và vùng xuyên màng. Theo các phương án khác, các trình tự nội bào và xuyên màng, ví dụ trình tự axit amin từ vị trí (hoặc vị trí tương đương của) 523, 524, 525, 526, 527, 528, 529, hoặc 530 của vùng HA2 đến đầu tận cùng C của vùng HA2

(đánh số theo trình tự SEQ ID NO: 1) được loại bỏ để tạo ra polypeptit hòa tan sau khi biểu hiện trong tế bào. Theo một số phương án, polypeptit được làm ổn định hơn nữa bằng cách đưa trình tự đã biết để tạo cấu trúc trime, tức là AYVRKDGEWVLL (SEQ ID NO: 80), tùy ý được nối thông qua cầu nối. Cầu nối này có thể tùy ý chứa vị trí phân cắt để xử lý sau đó theo quy trình mà người có hiểu biết trung bình trong lĩnh vực đã biết. Để dễ dàng tinh chế dạng hòa tan có thể bổ sung thêm trình tự đuôi, ví dụ đuôi his (HHHHHHH) được nối thông qua cầu nối ngắn, ví dụ EGR. Theo một số phương án cầu nối và trình tự đuôi his được thêm vào không có trình tự cuộn gấp được thể hiện. Theo một số phương án, trình tự nội bào và xuyên màng được thay thế bằng trình tự axit amin AGRHHHHHHH (SEQ ID NO: 97) hoặc SGRSLVPRGSPGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHH (SEQ ID NO: 82).

Các tác giả sáng chế đã xác định trước đó các kháng thể trung hòa phổ rộng được phân lập từ tế bào B nguyên thủy của người từ những người đã được chủng ngừa, một vài trong số đó đặc hiệu đối với virus cúm nhóm 1 (ví dụ CR6261, như được mô tả trong WO 2008/028946) và một vài trong số đó đặc hiệu đối với virus cúm nhóm 2 (ví dụ CR8020 như được mô tả trong WO 2010/130636). Phân tích chi tiết các epitop của các kháng thể đơn dòng này đã cho thấy nguyên nhân của việc không có tính phản ứng chéo của các kháng thể đặc hiệu này. Trong cả hai trường hợp sự có mặt của các glycan trong các phân tử HA nhóm 1 và nhóm 2 ở các vị trí khác nhau ít nhất đã giải thích một phần rằng các kháng thể là đặc hiệu nhóm. Với việc nhận diện các kháng thể tương tự CR9114 mà phản ứng chéo với nhiều phân tử HA nhóm 1 và nhóm 2, như được mô tả dưới đây, đã rõ ràng rằng hệ miễn dịch của người có thể kích hoạt các kháng thể trung hòa phổ rộng kháng virus cúm. Tuy nhiên, nhu cầu được đưa ra đối với lịch chủng ngừa hàng năm các kháng thể này hầu như không hoặc chỉ được kích hoạt ở mức độ rất thấp sau khi nhiễm hoặc chủng ngừa bằng virus cúm (mùa) thuộc các typ phụ H1 và/hoặc H3. Theo một số phương án, sáng chế do đó đề xuất các polypeptit trình diện vùng gốc của HA theo cách chính xác về cấu hình sao cho các epitop kích thích các kháng thể trung hòa phổ rộng được trình diện cho hệ miễn dịch khi không có vùng biến đổi trội miễn dịch. Bởi vì biết rằng mẫu hình glycan khác nhau giữa HA của H1 và của H3, và sự khác nhau này có thể dẫn đến đáp ứng của kháng thể bị giới hạn ở nhóm nhiều hơn, theo các phương án khác polypeptit theo sáng chế dựa trên các phân tử HA của nhóm 2 (ví dụ HA của H3). Như được thể hiện trong Ví dụ 3, dưới đây, khả năng trung hòa *in vitro* của CR9114 ở các typ phụ H1 cao hơn so với các typ phụ H3. Do đó, giả thuyết rằng epitop của CR9114 có khả năng truy

nhập hơn ở H1 so với các phân tử HA của H3, điều này có thể là do glycan ở N38 trong HA1 là chung nhau ở nhiều typ phụ HA của nhóm 2. Mong muốn không bị giới hạn bởi lý thuyết này, điều này có thể được giải thích rằng nếu polypeptit theo sáng chế dựa trên H1, các kháng thể tạo ra có nhiều khả năng bị cản trở bởi glycan ở vị trí N38 trên các phân tử HA của nhóm 2 và do đó phần nào ít hoạt động ở virus cúm nhóm 2. Do đó, để cho phép kích hoạt các kháng thể trung hòa phổ rộng mà hoạt động trên cả virus cúm nhóm 1 và nhóm 2 với hoạt tính tốt, theo một số phương án polypeptit vùng gốc theo sáng chế dựa trên HA của các typ phụ H3.

Người thường bị nhiễm virus cúm mùa chứa HA của typ phụ H1 hoặc H3. Mặc dù phơi nhiễm các virus cúm này, nhưng dường như các kháng thể trung hòa phổ rộng thường không xuất hiện trong trường hợp tự nhiên. Một trong những lý do của hiện tượng này là, ngoài sự có mặt của vùng đầu biến đổi trong HA, có thể là sự phơi nhiễm với một typ phụ mới có quan hệ gần gũi với typ phụ đã được phát hiện trước đó bằng cách này hay cách khác gây ra phổ đáp ứng ít hơn. Do đó có thể được ưu tiên hơn là phơi nhiễm của cá thể với trình tự typ phụ không liên quan. Do đó, theo phương án khác nữa, polypeptit vùng gốc theo sáng chế dựa trên HA của typ phụ nhóm 2 mà chứa asparagin (N) ở vị trí 38 trong HA1 (N38), và không phải là typ phụ H3.

Theo một số phương án, polypeptit dựa trên typ phụ virus cúm A. Theo một số phương án, polypeptit không dựa trên HA của H7.

Như được mô tả trên đây, polypeptit theo sáng chế được thiết kế không chỉ dựa trên các trình tự HA bố mẹ từ các typ phụ của virus cúm nhóm 1 làm vaccin cúm (ví dụ như H1 và H5), nhưng cũng có thể dựa trên các trình tự HA của các typ phụ virus cúm từ nhóm 2, cụ thể là các typ phụ của virus cúm nhóm 2 mà được sử dụng làm vaccin cúm, như H3. Theo sáng chế, các polypeptit được tạo cấu trúc mà bảo tồn epitop của CR8020 và CR8043 do các kháng thể này có khả năng trung hòa phổ rộng các chủng thuộc nhóm 2 (WO2010/130636). Trong các polypeptit này, dải beta ở phần dưới cùng của vùng gốc và vùng bao quanh nó cần được bảo tồn bởi vì vùng này là CR8020 và CR8043 liên kết với HA của H3.

Theo một số phương án, các vùng của HA là của typ phụ H3, tốt hơn của chủng A/Wisconsin/67/2005 (SEQ ID NO: 89), hoặc A/Hong Kong/1/1968 (SEQ ID NO: 121). Người có hiểu biết trung bình trong lĩnh vực cần phải hiểu rằng virus cúm khác A chứa HA của typ phụ H3 cũng có thể được sử dụng theo sáng chế.

Theo một số phương án, polypeptit chứa, hoặc bao gồm, đoạn polypeptit đầu tận cùng N của HA1 chứa axit amin từ vị trí 1 đến vị trí x của vùng HA1 của H3, tốt hơn là axit amin từ vị trí p đến vị trí x của vùng HA1, trong đó x là axit amin bất kỳ

nằm giữa các vị trí 56 và 69, như 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67 hoặc 68, của HA1 của H3, tốt hơn trong đó x là 61, 62, 63 hoặc 68. Theo một số phương án, đoạn polypeptit đầu tận cùng C của HA1 chứa axit amin từ vị trí y đến và bao gồm axit amin đầu tận cùng C của vùng HA1 của H3, trong đó y là axit amin bất kỳ nằm giữa và bao gồm các vị trí 292 và 325, của HA1 của H3, tốt hơn trong đó y là 293, 306, 318 hoặc 323.

Theo một số phương án, các vùng của HA là của typ phụ H3, tốt hơn là chủng A/Wisconsin/67/2005 (SEQ ID NO: 89), hoặc A/Hong Kong/1/1968 (SEQ ID NO: 121).

Theo sáng chế, vùng đầu được loại bỏ bằng cách loại bỏ một phần lớn của trình tự HA1 và nối lại các trình tự đầu tận cùng N và C thông qua một cầu nối ngắn. Đoạn bị loại bỏ có thể thay đổi về chiều dài, nhưng được ưu tiên là gốc cuối cùng của trình tự đầu tận cùng N của HA1 và gốc đầu tiên của trình tự đầu tận cùng C gần nhau về mặt không gian để tránh việc đưa chúng vào thông qua trình tự nối. Trong trình tự H3 các đoạn được loại bỏ có thể được đưa vào ở (các vị trí tương đương của) S62-P322, S63-P305 và T64-T317. Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp ví dụ như Clustal hoặc Muscle. Các phần còn lại của trình tự có thể được nối trực tiếp hoặc theo cách khác có thể đưa vào cầu nối linh động. Trình tự cầu nối có thể có chiều dài từ 1 đến 50 axit amin. Được ưu tiên là các cầu nối linh động có chiều dài giới hạn (nhỏ hơn hoặc bằng 10 axit amin), ví dụ GGG, GGGG, GSA, GSAG, GSAGSA, GSAGSAG hoặc tương tự. Chiều dài của đoạn bị loại bỏ cũng có thể thay đổi, ví dụ bằng cách giảm số lượng gốc trong đoạn được loại bỏ bằng cách bắt đầu ở (tương đương của) vị trí 63, 64, 65, 66, 67, hoặc làm tăng chiều dài của đoạn bị loại bỏ, bằng cách cắt ở vị trí 57, 58, 59, 60 hoặc 61. Tương tự, axit amin cuối cùng bị loại bỏ có thể ở (tương đương của) vị trí 317, 318, 319, 320 hoặc 321, hoặc làm tăng chiều dài của đoạn bị loại bỏ ở (tương đương của) vị trí 323, 324, 325, 326, hoặc 327. Quan trọng là nhận diện được sự thay đổi về chiều dài của đoạn được loại bỏ có thể ở trong phần được bù bằng cách làm cho phù hợp chiều dài của trình tự cầu nối, tức là đoạn được loại bỏ lớn hơn có thể được làm cho phù hợp với cầu nối dài hơn và ngược lại. Các polypeptit này cũng được bao gồm trong sáng chế.

Theo một số phương án, x là 61 và y là 323.

Theo một số phương án, x là 62 và y là 306.

Theo một số phương án, x là 63 và y là 318.

Theo một số phương án, x là (tương đương của) vị trí 62, 63, 64, 65, 66, hoặc vị trí 56, 57, 58, 59 hoặc 60.

Theo một số phương án, y là (tương đương của) vị trí 306, 318, 319, 320, 321 hoặc 322, hoặc (tương đương của) vị trí 324, 325, 326, 327, hoặc 328.

Theo một phương án, trình tự axit amin nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD tương ứng với trình tự axit amin giữa axit amin ở vị trí 400 và axit amin ở vị trí 420 của HA2 của trình tự SEQ ID NO: 89, hoặc các gốc axit amin ở các vị trí tương đương trong các chủng virus H3 khác, trong đó polypeptit chứa một hoặc nhiều đột biến trong trình tự axit amin nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD, tức là trình tự axit amin kéo dài từ axit amin 400-420 của trình tự SEQ ID NO: 89, hoặc các gốc axit amin tương đương trong các chủng virus cúm H3 khác.

Theo một số phương án, trình tự axit amin nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD của HA của virus cúm thuộc typ huyết thanh H3 chứa trình tự axit amin trong trình tự SEQ ID NO: 104.

Theo sáng chế, polypeptit chứa một hoặc nhiều đột biến trong trình tự axit amin nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD. Theo một số phương án, polypeptit chứa một hoặc nhiều đột biến trong Bảng 8, hoặc đột biến tương đương trong các chủng virus cúm khác thuộc typ phụ H3.

Theo sáng chế, vị trí cắt giữa HA1 và HA2 được loại bỏ. Theo một số phương án, việc loại bỏ vị trí cắt ở vị trí 345 (đánh số đề cập đến trình tự SEQ ID NO: 89) được làm đột biến (R345Q) để ngăn chặn sự tạo thành HA1 và HA2 từ HA0. Một cách tùy ý các gốc 347 đến 351 (IFGAI, một phần của peptit dung hợp) có thể được loại bỏ thêm để làm giảm đến mức thấp nhất sự tiếp xúc của các gốc kỵ nước với dung môi chứa nước. Điện tích dương ở vị trí cắt được bảo tồn 100% trong H3 và do đó đột biến này có thể được áp dụng trong tất cả các trình tự.

Việc loại bỏ vùng đầu ra khỏi vòng B giữa các gốc 400 đến 420 hiện được cho tiếp xúc với dung môi chứa nước. Trong các HA của H3 vòng này có tính bảo thủ cao (xem Bảng 9). Trình tự liên ứng là: 401 I(E/G)KTNEKFHQIEKEFSEVEGR 421 (SEQ ID NO: 104; đánh số đề cập đến trình tự SEQ ID NO: 89). Để làm tăng độ hòa tan của vòng này đối với polypeptit theo sáng chế trong cấu hình trước dung hợp và làm mất ổn định cấu hình sau khi dung hợp một số gốc kỵ nước phải được biến đổi thành gốc phân cực (S,T,N,Q), axit amin tích điện (R,H,K,D,E), hoặc phải làm tăng độ linh động bằng cách gây đột biến thành G. Gây đột biến đặc hiệu ở các vị trí 401, 408, 411,

415, 418, (đánh số đề cập đến trình tự SEQ ID NO: 89) sẽ góp phần làm ổn định polypeptit theo sáng chế.

Để làm ổn định cấu hình trước dung hợp của polypeptit theo sáng chế liên kết cộng hóa trị giữa hai phần xa nhau trong các trình tự bậc một nhưng gần nhau trong cấu hình được cuộn gấp trước dung hợp được đưa vào. Cuối cùng cầu disulfit có thể được thiết kế trong polypeptit theo sáng chế, tốt hơn là giữa (tương đương của) vị trí 326 và 438 trong chủng H3 A/Wisconsin/67/2005 (SEQ ID NO: 89). Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp như Clustal, Muscle etc. Cầu disulfit được thiết kế được tạo ra bằng cách gây đột biến ít nhất một (nếu gốc còn lại là xystein rồi), nhưng thường là hai gốc gần nhau về không gian với xystein, mà sẽ tạo thành liên kết cộng hóa trị một cách tự phát hoặc bằng phản ứng oxi hóa hoạt hóa giữa các nguyên tử lưu huỳnh của các gốc này. Cầu xystein thay thế có thể được tạo ra giữa (tương đương của) vị trí 334 và 393 trong chủng H3 A/Wisconsin/67/2005 (SEQ ID NO: 89) bằng cách gây đột biến các gốc này thành xystein. Trong một số trường hợp xystein ở (tương đương của) vị trí 321 được cải biến thành glyxin để tránh sự tạo thành cầu disulfit không mong muốn.

Theo một số phương án, polypeptit chứa một hoặc nhiều đột biến trong số các đột biến sau: F408S, I411T, F415S, V418G, I401R, K326C, S438C, T334C, I393C, C321G.

HA tự nhiên tồn tại ở dạng trime trên bề mặt tế bào. Hầu hết các tương tác giữa từng monome mà giữ dạng trime cùng nằm trong vùng đầu. Sau khi loại bỏ vùng đầu, do đó cấu trúc bậc ba được làm mất ổn định và vì vậy củng cố các tương tác giữa các monome trong phân tử bị cắt đứt sẽ làm tăng độ ổn định. Trong vùng gốc, trime hóa được trung gian bởi sự hình thành của môtip cuộn xoắn dạng trime. Bằng cách củng cố môtip này có thể tạo ra dạng trime thích hợp hơn. Trình tự liên ứng để tạo ra dạng trime xoắn cuộn, IEAIEKKIEAIEKKIEAIEKK, được đưa vào ở (tương đương của) vị trí 421 đến 441. Để tránh sự cản trở hình thành cầu disulfit giữa các vị trí 326 và 438 một trình tự thay thế ngắn hơn IEAIEKKIEAIEKKI ở (tương đương của) các vị trí 421 đến 435 cũng được sử dụng. Một cách khác là đưa vào trình tự RMKQIEDKIEEIESKQKKIEN, có nguồn gốc từ GCN4 và được biết là trime hóa, ở vị trí 421-441 hoặc trình tự ngắn hơn RMKQIEDKIEEIESK ở vị trí 421 đến 435.

Polypeptit theo sáng chế có thể chứa trình tự nội bào và vùng xuyên màng của HA vì vậy các polypeptit thu được trình diện trên bề mặt tế bào khi được biểu hiện trong tế bào. Theo các phương án khác, trình tự trong tế bào chất và trình tự xuyên

màng từ (tương đương của) vị trí 522 đến đầu tận cùng C được loại bỏ do đó polypeptit được tiết ra (hòa tan được) được sản sinh sau khi biểu hiện trong tế bào. Tùy ý, một số gốc bổ sung có thể được bao gồm trong protein hòa tan được bằng cách loại bỏ trình tự từ (tương đương của) 523, 524, 525, 526, 527, 528 hoặc 529. Polypeptit hòa tan được có thể được làm ổn định hơn nữa bằng cách đưa vào trình tự đã biết để tạo cấu trúc trimer, tức là AYVRKDGEWVLL (SEQ ID NO: 143)(trình tự 'cuộn gấp'), tùy ý được nối thông qua cầu nối. Cầu nối này có thể tùy ý chứa vị trí phân cắt để xử lý sau đó theo quy trình mà người có hiểu biết trung bình trong lĩnh vực đã biết. Để dễ dàng tinh chế dạng hòa tan có thể bổ sung thêm trình tự đuôi, ví dụ đuôi his (HHHHHHH) được nối thông qua cầu nối ngắn, ví dụ EGR. Theo một số phương án cầu nối và trình tự đuôi his được thêm vào không có trình tự cuộn gấp được thể hiện.

Theo sáng chế, trình tự axit amin từ vị trí 530 (đánh số theo trình tự SEQ ID NO: 1) đến gốc axit amin đầu tận cùng C của vùng HA2 có thể được loại bỏ và thay thế bằng các trình tự sau: EGRHHHHHHH (SEQ ID NO: 81), hoặc SGRSLVPRGSPGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHH (SEQ ID NO: 82).

Theo một số phương án đoạn gốc đầu tận cùng N của HA1 không chứa trình tự tín hiệu. Người có hiểu biết trung bình trong lĩnh vực cần phải hiểu rằng trình tự dẫn đầu (hoặc trình tự tín hiệu) mà điều khiển sự vận chuyển protein trong quá trình sản sinh (ví dụ tương ứng với axit amin 1-17 của trình tự SEQ ID NO: 89), thường sẽ không có mặt trong polypeptit cuối cùng, ví dụ được sử dụng trong vacxin. Do đó, theo một số phương án, polypeptit theo sáng chế bao gồm trình tự axit amin không có trình tự dẫn đầu.

Theo sáng chế, polypeptit không dựa trên các phân tử HA của virus cúm B. Chủng virus cúm typ B gây bệnh cúm lạnh tính ở người. Sự biến đổi kháng nguyên trong HA trong chủng virus cúm typ B nhỏ hơn sự biến đổi quan sát được trong chủng virus cúm A. Hai giống khác nhau về di truyền và kháng nguyên của virus cúm B đang lưu hành ở người, được đại diện bởi giống B/Yamagata/16/88 (còn được gọi là B/Yamagata) và B/Victoria/2/87 (B/Victoria) (Ferguson et al., 2003). Mặc dù phổ gây bệnh do virus cúm B thường nhẹ hơn bệnh do virus cúm A gây ra, nhưng vẫn thường quan sát được ca bệnh trầm trọng cần phải nằm viện do nhiễm virus cúm B.

Theo sáng chế các polypeptit được đề xuất mà mô phỏng các epitop đặc hiệu của CR6261 và CR9114, và có thể được sử dụng làm các polypeptit gây miễn dịch, ví dụ để kích hoạt các kháng thể trung hòa chéo khi được sử dụng in vivo, riêng hoặc kết

hợp với chất điều trị dự phòng và/hoặc điều trị chữa bệnh khác. Với “các kháng thể trung hòa chéo”, các kháng thể đề chỉ rằng có khả năng trung hòa ít nhất là hai, tốt hơn ít nhất là ba, bốn, hoặc năm typ phụ khác nhau của virus cúm A thuộc nhóm phát sinh chủng loại 1, và/hoặc ít nhất là hai, tốt hơn ít nhất là ba, bốn, hoặc năm typ phụ khác nhau của virus cúm A thuộc nhóm phát sinh chủng loại 2, và/hoặc ít nhất là hai, typ phụ khác nhau của virus cúm B, cụ thể là ít nhất là tất cả các chủng virus mà được trung hòa bởi CR6261 và CR9114.

Polypeptit theo sáng chế không chứa HA1 chiều dài đủ. Theo một số phương án, các polypeptit gây miễn dịch hầu như nhỏ hơn HA0, tốt hơn không có toàn bộ hoặc hầu như toàn bộ vùng đầu hình cầu của HA. Tốt hơn, các polypeptit gây miễn dịch có chiều dài không nhiều hơn 360, tốt hơn không nhiều hơn 350, 340, 330, 320, 310, 305, 300, 295, 290, 285, 280, 275, hoặc 270 axit amin. Theo một phương án, polypeptit gây miễn dịch có chiều dài từ 250 đến 350, tốt hơn là từ 260 đến 340, tốt hơn là từ 270 đến 330, tốt hơn là từ 270 đến 330 axit amin.

Theo một số phương án, polypeptit liên kết có chọn lọc với kháng thể CR6261 và/hoặc CR9114. Theo một phương án, polypeptit không liên kết với kháng thể CR8057. Theo một phương án, CR6261 chứa vùng biến đổi của chuỗi nặng chứa trình tự axit amin nêu trong SEQ ID NO: 20 và vùng biến đổi của chuỗi nhẹ chứa trình tự axit amin nêu trong SEQ ID NO: 21; CR9114 chứa vùng biến đổi của chuỗi nặng chứa trình tự axit amin nêu trong SEQ ID NO: 18 và vùng biến đổi của chuỗi nhẹ chứa trình tự axit amin nêu trong SEQ ID NO: 19. Theo một phương án, CR8057 chứa vùng biến đổi của chuỗi nặng chứa trình tự axit amin nêu trong SEQ ID NO: 22 và vùng biến đổi của chuỗi nhẹ chứa trình tự axit amin nêu trong SEQ ID NO: 23.

Như được mô tả trên đây, polypeptit chứa vùng HA1 của hemagglutinin của virus cúm mà bao gồm đoạn gốc đầu tận cùng N của HA1 được liên kết cộng hóa trị bằng trình tự nối gồm 0- 50 các gốc axit amin với đoạn gốc đầu tận cùng C của HA1. Trình tự nối không có trong tự nhiên, hoặc kiểu đại, HA. Theo một số phương án, cầu nối là peptit mà bao gồm một gốc axit amin, hai gốc axit amin hoặc ít hơn, ba gốc axit amin hoặc ít hơn, bốn gốc axit amin hoặc ít hơn, năm gốc axit amin hoặc ít hơn, mười gốc axit amin hoặc ít hơn, 15 gốc axit amin hoặc ít hơn, hoặc 20 gốc axit amin hoặc ít hơn hoặc 30 gốc axit amin hoặc ít hơn hoặc 40 gốc axit amin hoặc ít hơn hoặc 50 gốc axit amin hoặc ít hơn. Theo phương án cụ thể, trình tự nối là trình tự được chọn từ nhóm bao gồm G, GS, GGG, GSG, GSA, GSGS, GSAG, GGGG, GSAGS, GSGSG, GSAGSA, GSAGSAG, và GSGSGSG.

Sáng chế còn đề cập đến các phương pháp tạo ra polypeptit theo sáng chế, cụ thể là đề cập đến trình tự axit amin của polypeptit vùng gốc của HA theo sáng chế, cũng như polypeptit có thể thu được hoặc thu được bằng các phương pháp này. Theo một số phương án, phương pháp bao gồm các bước sau:

- Cung cấp trình tự axit amin HA0 của virus cúm, ví dụ trình tự HA0 của virus cúm thuộc typ huyết thanh H1, H5 hoặc H3;

- Loại bỏ vị trí cắt giữa HA1 và HA2, tốt hơn là bằng cách gây đột biến axit amin đầu tận cùng C của HA1 thành axit amin khác với arginin (R) hoặc lysin (K);

- Loại bỏ trình tự axit amin của vùng đầu hình cầu từ trình tự HA0; Bước này được thực hiện bằng cách loại bỏ một đoạn của vùng HA1 giữa axit amin ở vị trí x và axit amin ở vị trí y, và nối lại đoạn đầu tận cùng N (kéo dài từ axit amin ở vị trí 1 đến và bao gồm axit amin ở vị trí x của HA1) và đoạn đầu tận cùng C của HA1 (kéo dài từ axit amin y đến axit amin đầu tận cùng C của HA1), do đó thu được, tùy ý thông qua trình tự nối gồm 0-50 axit amin.

- Làm tăng độ ổn định của cấu hình trước khi dung hợp và loại bỏ việc làm ổn định cấu hình sau khi dung hợp của HA cải biến, tốt hơn là bằng cách đưa một hoặc nhiều đột biến vào trong trình tự axit amin nối gốc ở đầu tận cùng C của chuỗi xoắn A với gốc ở đầu tận cùng N của chuỗi xoắn CD, tốt hơn là trong trình tự axit amin kéo dài từ axit amin 402-418 của HA của H1, cụ thể là chứa trình tự axit amin MNTQFTAVGKEFN(H/K)LE(K/R) (SEQ ID NO: 17) hoặc I(E/G)KTNEKFHQIEKEFSEVEGR 421 (SEQ ID NO: 104) đối với HA của H3. Đột biến tốt hơn là đột biến thay thế các gốc axit amin kỵ nước thành các gốc axit amin ưa nước.

- Đưa một hoặc nhiều cầu disulfit vào polypeptit vùng gốc của HA.

Theo sáng chế, việc loại bỏ vị trí cắt giữa HA1 và HA2 có thể đạt được bằng cách gây đột biến R (một số ít trường hợp là K) thành Q ở vị trí P1 (xem ví dụ Sun et al, 2010 để biết giải thích về danh pháp của vị trí cắt (vị trí 343 trong trình tự SEQ ID NO: 1). Việc gây đột biến thành Q được ưu tiên nhưng S, T, N, D hoặc E là các dạng thay thế.

Việc loại bỏ vùng đầu có thể đạt được, ví dụ bằng cách loại bỏ axit amin 53 đến 320 từ SEQ ID NO: 1. Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp ví dụ như Clustal hoặc Muscle. Các phần còn lại của trình tự có thể được nối trực tiếp, hoặc theo cách khác có thể đưa vào cầu nối linh động. Trình tự cầu nối có thể có chiều dài từ 1 đến 50 axit amin. Được ưu tiên là các

cầu nối linh động có chiều dài giới hạn (nhỏ hơn hoặc bằng 10 axit amin), ví dụ GGG, GGGG, GSA, GSAG, GSAGSA, GSAGSAG hoặc tương tự. Chiều dài của đoạn bị loại bỏ cũng có thể thay đổi, ví dụ bằng cách bắt đầu loại bỏ ở (tương đương của) vị trí (x), ví dụ ở vị trí 54, 55, 56, 57 hoặc 58, hoặc làm tăng chiều dài của đoạn bị loại bỏ, bằng cách cắt ở vị trí 47, 48, 49, 50, 51, hoặc 52. Tương tự, axit amin cuối cùng bị loại bỏ có thể ở (tương đương của) vị trí (y), như 315, 316, 317, 318 hoặc 319, hoặc làm tăng chiều dài của đoạn bị loại bỏ ở (tương đương của) vị trí 321, 322, 323, 324, hoặc 325. Quan trọng là nhận diện được sự thay đổi về chiều dài của đoạn được loại bỏ có thể ở trong phần được bù bằng cách làm cho phù hợp chiều dài của trình tự cầu nối, tức là đoạn được loại bỏ lớn hơn có thể được làm cho phù hợp với cầu nối dài hơn và ngược lại. Các polypeptit này cũng có thể được bao gồm trong sáng chế.

Theo sáng chế, độ hòa tan của vòng giữa chuỗi cuộn xoắn A và chuỗi cuộn xoắn CD được tăng lên. Vòng này được tạo thành bởi (tương đương của) các gốc 402 đến 418 trong chủng H1 A/Brisbane/59/2007 (SEQ ID NO: 1). Do đó, độ ổn định của cấu hình trước khi dung hợp được tăng lên và cấu hình sau khi dung hợp của HA được cải biến được làm mất ổn định. Vòng này có tính bảo thủ cao trong trình tự H1, như được thể hiện trong bảng 6 dưới đây. Điều này có thể đạt được chẳng hạn bằng cách thay thế axit amin I, L, F hoặc V trong vòng này bằng các bản sao ưa nước. Vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp ví dụ như Clustal hoặc Muscle. Đột biến thành glyxin làm mất ổn định cấu hình sau khi dung hợp do độ linh động cao của axit amin này dẫn đến sự giảm độ ổn định của chuỗi cuộn xoắn sau dung hợp được tạo thành bởi phần này của trình tự HA. Trình tự liên ứng mô tả vòng giữa gốc 402-418 của HA của virus cúm thuộc typ huyết thanh H1 là (SEQ ID NO: 17) MNTQFTAVGKEFN(H/K)LE(K/R). Trong một số polypeptit theo sáng chế axit amin ở các vị trí 406, 409, 413 và/hoặc 416 (hoặc vị trí tương đương của chúng, như được xác định từ sắp xếp và so sánh trình tự) là axit amin phân cực (S,T,N,Q), tích điện (R,H,K,D,E) hoặc linh động (G). Cũng có thể là tổ hợp của các đột biến này ở các vị trí này, ví dụ F406S, V409T, L416S như trong trình tự SEQ ID NO: 10 và SEQ ID NO: 14. Trong một số trường hợp đột biến để phục hồi axit amin liên ứng được ưu tiên, ví dụ trong đó V hoặc M ở vị trí 404 (thành T), V ở 408 (thành A) hoặc 410 (thành G) hoặc I ở 414 (thành N); tỷ lệ đột biến trong trình tự ở các axit amin này là rất thấp. Tổng quát về đột biến được mô tả trên đây mà xác định đặc tính polypeptit theo sáng chế được đưa ra trong các Bảng 6.

Theo sáng chế, một hoặc nhiều cầu disulfidit được đưa vào trong polypeptit vùng gốc, tốt hơn là giữa axit amin ở (hoặc vị trí tương đương của) vị trí 324 và 436 trong chủng H1 A/Brisbane/59/2007: SEQ ID NOs: 13-16. Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp như Clustal, Muscle etc. Cầu disulfidit được thiết kế được tạo ra bằng cách gây đột biến ít nhất một (nếu gốc còn lại là xystein rồi), nhưng thường là hai gốc gần nhau về không gian với xystein, mà sẽ tạo thành liên kết cộng hóa trị một cách tự phát hoặc bằng phản ứng oxi hóa hoạt hóa giữa các nguyên tử lưu huỳnh của các gốc này.

Các polypeptit có thể thu được bằng phương pháp này cũng là một phần của sáng chế.

HA tự nhiên tồn tại ở dạng trime trên bề mặt tế bào. Hầu hết các tương tác giữa từng monome mà giữ dạng trime cùng nằm trong vùng đầu. Sau khi loại bỏ vùng đầu, cấu trúc bậc ba do đó được làm mất ổn định và vì vậy củng cố các tương tác giữa các monome trong phân tử bị cắt cụt sẽ làm tăng độ ổn định. Trong vùng gốc, trime hóa được trung gian bởi sự hình thành của môtip cuộn xoắn dạng trime. Bằng cách củng cố môtip này có thể tạo ra dạng trime thích hợp hơn. Theo sáng chế, trình tự liên ứng để tạo ra dạng trime xoắn cuộn, IEAIEKKIEAIEKKIE (SEQ ID NO: 83), có thể được đưa vào trong polypeptit theo sáng chế ở (tương đương của) vị trí 418 đến 433. Theo một số phương án, trình tự MKQIEDKIEEIESKQ (SEQ ID NO: 84), có nguồn gốc từ GCN4 và được biết là trime hóa được đưa vào ở (tương đương của) vị trí 419-43. Theo một số phương án, mặt trung gian của trime được làm ổn định bằng cách cải biến M420, L423, V427, G430 thành isoleuxin.

Theo một số phương án, polypeptit bao gồm trình tự axit amin được chọn từ nhóm bao gồm SEQ ID NO: 3-16, SEQ ID NO: 44-53, SEQ ID NO: 111-114, SEQ ID NO: 119-120, SEQ ID NO: 125, 126, 130, SEQ ID NO: 144-175 và SEQ ID NO: 177-187.

Theo một số phương án, polypeptit được chọn từ nhóm bao gồm SEQ ID NO: 45, SEQ ID NO: 113 và SEQ ID NO: 130.

Người có hiểu biết trung bình trong lĩnh vực cần phải hiểu rằng trình tự dẫn đầu (hoặc trình tự tín hiệu) mà điều khiển sự vận chuyển protein trong quá trình sản sinh (ví dụ tương ứng với axit amin 1-17 của trình tự SEQ ID NO: 1), không có mặt trong polypeptit cuối cùng, ví dụ được sử dụng trong vacxin. Theo một số phương án, polypeptit theo sáng chế do đó bao gồm trình tự axit amin không có trình tự dẫn đầu.

Các polypeptit vùng gốc hemagglutinin của virus cúm này có thể được điều chế theo kỹ thuật bất kỳ được xem là phù hợp với người có hiểu biết trung bình trong lĩnh vực, bao gồm các kỹ thuật mô tả dưới đây.

Do đó, các polypeptit gây miễn dịch theo sáng chế có thể được tổng hợp dưới dạng trình tự ADN bằng các phương pháp chuẩn đã biết trong lĩnh vực và được tách dòng và sau đó được biểu hiện, *in vitro hoặc in vivo*, sử dụng enzym giới hạn và phương pháp phù hợp đã biết trong lĩnh vực. Do đó sáng chế cũng đề cập đến các phân tử axit nucleic mã hóa các polypeptit được mô tả trên đây. Sáng chế còn đề cập đến các vector chứa axit nucleic mã hóa polypeptit theo sáng chế. Theo một số phương án, phân tử axit nucleic theo sáng chế là một phần của vector, ví dụ plasmit. Các vector này có thể được thao tác dễ dàng bằng các phương pháp mà người có hiểu biết trung bình trong lĩnh vực đã biết, và có thể thể ví dụ được thiết kế để có khả năng sao chép trong tế bào nhân sơ và/hoặc tế bào nhân thật. Ngoài ra, nhiều vector có thể trực tiếp hoặc ở dạng phân đoạn mong muốn được phân lập từ đó được sử dụng để biến đổi tế bào nhân thật và sẽ tích hợp toàn bộ hoặc một phần trong bộ gen của các tế bào này, tạo ra tế bào chủ thích hợp chứa axit nucleic mong muốn trong bộ gen của chúng. Vector được sử dụng có thể là vector bất kỳ mà thích hợp để tách dòng ADN và có thể được sử dụng để phiên mã axit nucleic quan tâm. Khi tế bào chủ được sử dụng, vector được ưu tiên là vector tích hợp. Theo cách khác, vector này có thể là vector sao chép thể bổ sung.

Người có hiểu biết trung bình trong lĩnh vực có khả năng lựa chọn các vector biểu hiện phù hợp, và chèn các trình tự axit nucleic theo sáng chế theo cách có chức năng. Để thu được sự biểu hiện của các trình tự axit nucleic mã hóa polypeptit, người có hiểu biết trung bình trong lĩnh vực đã biết rằng các trình tự có khả năng điều khiển sự biểu hiện có thể được nối theo cách có chức năng với các trình tự axit nucleic mã hóa polypeptit, tạo thành các phân tử axit nucleic tái tổ hợp mã hóa protein hoặc polypeptit ở dạng biểu hiện được. Nhìn chung, trình tự khởi đầu phiên mã được đặt phía trước các trình tự cần được biểu hiện. Nhiều vector biểu hiện có sẵn trong lĩnh vực này, ví dụ hàng loạt vector pcADN và pEF của hãng Invitrogen, pMSCV và pTK-Hyg của hãng BD Sciences, pCMV-Script của hãng Stratagene, v.v., có thể được sử dụng để thu được các trình tự khởi đầu và/hoặc kết thúc phiên mã phù hợp, trình tự polyA, và tương tự. Trong đó trình tự mã hóa polypeptit quan tâm được chèn vào thích hợp có dẫn đến các trình tự điều khiển quá trình phiên mã và dịch mã của polypeptit được mã hóa, cát xét biểu hiện thu được có tác dụng sản sinh ra polypeptit quan tâm, được gọi là sự biểu hiện. Các trình tự điều khiển sự biểu hiện có thể bao gồm trình tự khởi đầu phiên mã, trình tự tăng cường và tương tự, và sự kết hợp của

chúng. Các trình tự này có khả năng hoạt động trong tế bào chủ, qua đó điều khiển sự biểu hiện của các trình tự axit nucleic mà được nối theo cách có chức năng với chúng. Người có hiểu biết trung bình trong lĩnh vực nhận thấy rằng các trình tự khởi đầu phiên mã khác nhau có thể được sử dụng để thu được sự biểu hiện của gen trong tế bào chủ. Các trình tự khởi đầu phiên mã có thể cấu thành hoặc được điều hòa, và có thể thu được từ các nguồn khác nhau, bao gồm nguồn virus, nhân sơ hoặc nhân thật, hoặc được thiết kế nhân tạo. Sự biểu hiện của các axit nucleic quan tâm có thể là từ trình tự khởi đầu phiên mã tự nhiên hoặc dẫn xuất của nó hoặc từ trình tự khởi đầu phiên mã hoàn toàn khác loài (Kaufman, 2000). Một số trình tự khởi đầu phiên mã đã biết và được sử dụng nhiều để biểu hiện trong tế bào nhân thật chứa các trình tự khởi đầu phiên mã có nguồn gốc từ virus, như adenovirus, ví dụ trình tự khởi đầu phiên mã của E1A, các trình tự khởi đầu phiên mã có nguồn gốc từ cytomegalovirus (CMV), như trình tự khởi đầu phiên mã từ CMV ngay ban đầu (immediate early-IE) (còn được gọi là trình tự khởi đầu phiên mã CMV) (có thể thu được chẳng hạn từ pcADN, Invitrogen), các trình tự khởi đầu phiên mã có nguồn gốc từ Virus Simian 40 (SV40) (Das et al, 1985), và tương tự. Các trình tự khởi đầu phiên mã thích hợp cũng có thể có nguồn gốc từ tế bào nhân thật, như các trình tự khởi đầu phiên mã methallothionein (MT), yếu tố kéo dài 1 α (EF-1 α) trình tự khởi đầu phiên mã (Gill et al., 2001), trình tự khởi đầu phiên mã ubiquitin C hoặc UB6 (Gill et al., 2001), trình tự khởi đầu phiên mã actin, trình tự khởi đầu phiên mã globulin miễn dịch, các trình tự khởi đầu phiên mã sốc nhiệt, và tương tự. Thử nghiệm về chức năng và chiều dài của trình tự khởi đầu phiên mã là vấn đề thông thường đối với người có hiểu biết trung bình trong lĩnh vực, và nhìn chung có thể chẳng hạn bao gồm tách dòng gen thử nghiệm như lacZ, luxiferaza, GFP, etc. nằm sau trình tự trình tự khởi đầu phiên mã, và thử nghiệm về sự biểu hiện của gen thử nghiệm. Tất nhiên, các trình tự khởi đầu phiên mã có thể được làm thay đổi bằng cách loại bỏ, thêm, gây đột biến về trình tự trong đó, và được thử nghiệm về chức năng, để tìm ra các trình tự khởi đầu phiên mã mới, được làm yếu hoặc được cải thiện. Theo sáng chế, việc chọn lựa các trình tự khởi đầu phiên mã mạnh tạo ra mức độ phiên mã cao trong tế bào nhân thật được ưu tiên.

Các cấu trúc này có thể được chuyển nhiễm vào tế bào nhân thật (ví dụ tế bào thực vật, nấm, nấm men hoặc động vật) hoặc hệ biểu hiện nhân sơ phù hợp như *E. coli* bằng các phương pháp mà người có hiểu biết trung bình trong lĩnh vực đã biết. Trong một số trường hợp trình tự “đuôi” thích hợp (ví dụ như, nhưng không chỉ giới hạn ở, đuôi his-, myc-, strep-, hoặc flag-) hoặc protein hoàn chỉnh (ví dụ như, nhưng không chỉ giới hạn ở, protein liên kết maltoza hoặc glutathion S transferaza) có thể được

thêm vào các trình tự theo sáng chế để cho phép tinh chế và/hoặc nhận diện các polypeptit từ tế bào hoặc dịch nổi. Tùy ý, trình tự chứa vị trí phân giải protein đặc hiệu có thể được bao gồm để sau đó loại bỏ đuôi bằng cách phân giải protein.

Các polypeptit tinh sạch có thể được phân tích bằng phương pháp quang phổ đã biết trong lĩnh vực (ví dụ phương pháp quang phổ lượng sắc vòng, phương pháp quang phổ hồng ngoại biến đổi Fourier và phương pháp quang phổ NMR hoặc tinh thể học tia X) để kiểm tra sự có mặt của cấu trúc mong muốn như chuỗi cuộn xoắn và dải beta. ELISA, Octet và FACS và phương pháp tương tự có thể được sử dụng để kiểm tra liên kết của polypeptit theo sáng chế với các kháng thể trung hòa phổ rộng được mô tả trước (CR6261, CR9114, CR8057). Do đó, các polypeptit theo sáng chế có cấu hình chính xác có thể được chọn lọc.

Sáng chế còn đề cập đến các chế phẩm gây miễn dịch chứa lượng có hiệu quả điều trị bệnh của ít nhất một polypeptit và/hoặc các axit nucleic theo sáng chế. Theo một số phương án, các chế phẩm chứa các polypeptit chứa các vùng gốc hemagglutinin từ (hoặc dựa trên) HA của một typ phụ virus cúm, ví dụ dựa trên HA của virus cúm chứa HA của ví dụ typ phụ H1 hoặc H7. Theo một số phương án, các chế phẩm chứa các polypeptit chứa các vùng gốc hemagglutinin dựa trên HA của hai hoặc nhiều typ phụ khác nhau của virus cúm, ví dụ các chế phẩm chứa cả các polypeptit chứa các vùng gốc hemagglutinin dựa trên HA của typ phụ H1 và các polypeptit chứa các vùng gốc hemagglutinin dựa trên HA của typ phụ H7.

Tốt hơn, các chế phẩm gây miễn dịch còn chứa chất mang được dụng. Trong ngữ cảnh này, thuật ngữ "được dụng" chỉ chất mang, với liều lượng và nồng độ phù hợp được sử dụng, sẽ không gây ra các tác dụng không mong muốn hoặc có hại cho đối tượng sử dụng các chất mang này. Các chất mang và tá dược được dụng này là đã biết trong lĩnh vực (xem Remington's Pharmaceutical Sciences, 18th edition, A. R. Gennaro, Ed., Mack Publishing Company [1990]; Pharmaceutical Formulation Development of Peptides và Proteins, S. Frokjaer và L. Hovgaard, Eds., Taylor & Francis [2000]; và Handbook of Pharmaceutical Excipients, 3rd edition, A. Kibbe, Ed., Pharmaceutical Press [2000]). Thuật ngữ "chất mang" chỉ chất pha loãng, chất hỗ trợ, tá dược, hoặc chất dẫn thuốc mà được sử dụng cùng với chế phẩm. Dung dịch nước muối và dung dịch nước dextroza và dung dịch glyxerol có thể ví dụ được sử dụng làm chất mang lỏng, cụ thể là làm các dung dịch tiêm được. Dạng bào chế chính xác cần phải phù hợp với đường sử dụng. Polypeptit và/hoặc các phân tử axit nucleic tốt hơn là được bào chế và sử dụng dưới dạng dung dịch vô trùng. Dung dịch vô trùng được điều chế bằng cách lọc vô trùng hoặc bằng các phương pháp khác mà người có hiểu biết

trung bình trong lĩnh vực đã biết. Dung dịch này sau đó có thể được làm đông khô hoặc được nạp vào các vật chứa phân liều dược phẩm. Độ pH của dung dịch thường nằm trong khoảng pH 3,0 đến 9,5, ví dụ pH 5,0 đến 7,5.

Sáng chế còn đề cập đến các phương pháp gây đáp ứng miễn dịch ở đối tượng, phương pháp này bao gồm sử dụng cho đối tượng, polypeptit, phân tử axit nucleic và/hoặc chế phẩm gây miễn dịch như được mô tả trên đây. Đối tượng theo sáng chế tốt hơn là động vật có vú mà có thể được gây nhiễm bằng tác nhân gây bệnh lây nhiễm, cụ thể là virus cúm, hoặc theo cách khác có thể có lợi từ việc gây đáp ứng miễn dịch, đối tượng này chẳng hạn là loài gặm nhấm, ví dụ chuột nhắt, chồn, hoặc vật nuôi trong nhà hoặc trang trại, hoặc động vật linh trưởng không phải người, hoặc người. Tốt hơn, đối tượng này là đối tượng người. Do đó sáng chế đề xuất các phương pháp gây đáp ứng miễn dịch đối với hemagglutinin (HA) của virus cúm, cụ thể là virus cúm A thuộc nhóm 1 và/hoặc nhóm 2, như virus cúm chứa HA thuộc typ phụ H1, H2, H3, H4, H5, H7 và/hoặc H10, và/hoặc của virus cúm B, ở đối tượng sử dụng polypeptit, các axit nucleic và/hoặc các chế phẩm gây miễn dịch được mô tả ở đây. Theo một số phương án, đáp ứng miễn dịch được gây ra có hiệu quả phòng ngừa và/hoặc điều trị nhiễm virus cúm do typ phụ của virus cúm A thuộc nhóm 1 và/hoặc nhóm 2 và/hoặc virus cúm B gây ra. Theo một số phương án, đáp ứng miễn dịch được gây ra bởi polypeptit, các axit nucleic và/hoặc các chế phẩm gây miễn dịch được mô tả ở đây có hiệu quả phòng ngừa và/hoặc điều trị nhiễm virus cúm A và/hoặc B gây ra bởi hai, ba, bốn, năm hoặc sáu typ phụ của viruts cúm A và/hoặc B.

Bởi vì đã biết rằng các protein và/hoặc các phân tử axit nucleic nhỏ luôn luôn không gây ra một cách hiệu quả đáp ứng miễn dịch có hiệu lực, có thể cần thiết để làm tăng tính gây miễn dịch của polypeptit và/hoặc các phân tử axit nucleic bằng cách bổ sung tá dược. Theo một số phương án, các chế phẩm gây miễn dịch được mô tả ở đây chứa, hoặc được sử dụng kết hợp với, tá dược. Tá dược để sử dụng kết hợp với chế phẩm được mô tả ở đây có thể được sử dụng trước, đồng thời, hoặc sau khi sử dụng chế phẩm này. Ví dụ về các tá dược thích hợp bao gồm các muối nhôm như nhôm hydroxit và/hoặc nhôm phosphat; các chế phẩm nhũ tương trong dầu (hoặc các chế phẩm dầu trong nước), bao gồm các nhũ tương squalen-nước, như MF59 (xem ví dụ WO 90/14837); chế phẩm saponin, ví dụ như QS21 và các phức hợp kích thích miễn dịch (ISCOMS) (xem ví dụ US 5,057,540; WO 90/03184, WO 96/11711, WO 2004/004762, WO 2005/002620); các chất dẫn xuất từ vi khuẩn hoặc vi sinh vật, ví dụ về các chất này là monophosphoryl lipid A (MPL), MPL 3-O-deaxylat hóa (3dMPL), CpG-môtip chứa oligonucleotit, độc tố vi khuẩn ADP-ribosyl hóa hoặc dạng đột biến

của nó, như nội độc tố LT không bền nhiệt từ *E. coli*, độc tố khuẩn tả CT, độc tố ho gà PT, hoặc độc tố uốn ván TT, M chất nền (Isconova). Ngoài ra, các kỹ thuật làm cho có khả năng miễn dịch đã biết có thể được sử dụng, như dung hợp polypeptit theo sáng chế với các protein đã biết trong lĩnh vực để tăng cường đáp ứng miễn dịch (ví dụ độc tố uốn ván, CRM197, rCTB, flagellin từ vi khuẩn hoặc các protein khác) hoặc bao gồm polypeptit trong virosom, hoặc kết hợp các kỹ thuật này. Các ví dụ không giới hạn khác có thể được sử dụng là các ví dụ được bộc lộ bởi Coffman et al. (2010).

Theo một phương án, các polypeptit vùng gốc hemagglutinin của virus cúm này theo sáng chế được kết hợp vào vectơ dạng hạt tương tự virus (VLP). VLP thường chứa (các) polypeptit của virus thông thường có nguồn gốc từ (các) polypeptit cấu trúc của virus. Tốt hơn, VLP không có khả năng sao chép. Theo một số phương án, VLP có thể không có bộ gen hoàn chỉnh của virus hoặc chứa một phần của bộ gen virus. Theo một số phương án, VLP không có khả năng gây nhiễm tế bào. Theo một số phương án, VLP biểu hiện trên bề mặt của chúng một hoặc nhiều gốc đích của virus (ví dụ, glycoprotein bề mặt của virus) hoặc không phải của virus (ví dụ, kháng thể hoặc protein) mà người có hiểu biết trung bình trong lĩnh vực đã biết.

Theo phương án cụ thể, polypeptit theo sáng chế được kết hợp vào virosom. Virosom chứa polypeptit theo sáng chế có thể được tạo ra bằng cách sử dụng các kỹ thuật mà người có hiểu biết trung bình trong lĩnh vực đã biết. Ví dụ, virosom có thể được tạo ra bằng cách phá vỡ virus tinh sạch, tách biệt bộ gen, và lắp ghép lại các hạt bằng các protein của virus (ví dụ, polypeptit vùng gốc hemagglutinin của virus cúm) và lipid để tạo ra các hạt lipid chứa các protein của virus.

Sáng chế còn đề cập đến các polypeptit được mô tả trên đây, các axit nucleic và/hoặc các chế phẩm gây miễn dịch để gây đáp ứng miễn dịch ở đối tượng kháng HA của virus cúm, cụ thể là được sử dụng làm vaccin. Các polypeptit vùng gốc hemagglutinin của virus cúm này, các axit nucleic mã hóa các polypeptit này, hoặc các vectơ chứa các axit nucleic này hoặc các polypeptit được mô tả ở đây do đó có thể được sử dụng để kích hoạt các kháng thể trung hòa kháng virus cúm, ví dụ, kháng vùng gốc của virus cúm hemagglutinin. Cụ thể, sáng chế đề cập đến các polypeptit, các axit nucleic, và/hoặc các chế phẩm gây miễn dịch như được mô tả trên đây được sử dụng làm vaccin trong việc phòng ngừa và/hoặc điều trị bệnh hoặc tình trạng bệnh lý gây ra bởi virus cúm A thuộc nhóm phát sinh chủng loại 1 và/hoặc nhóm phát sinh chủng loại 2 và/hoặc virus cúm B. Theo một phương án, vaccin có thể được sử dụng trong việc phòng ngừa và/hoặc điều trị các bệnh gây ra bởi hai, ba, bốn, năm, sáu hoặc nhiều hơn typ phụ khác nhau của nhóm phát sinh chủng loại 1 và/hoặc 2 và/hoặc virus

cúm B. Polypeptit theo sáng chế có thể được sử dụng sau tổng hợp *in vitro* hoặc trong hệ biểu hiện tế bào thích hợp, bao gồm tế bào vi khuẩn và tế bào nhân thật, hoặc theo cách khác, có thể được biểu hiện *in vivo* ở đối tượng cần, bằng cách biểu hiện axit nucleic mã hóa polypeptit gây miễn dịch. Vacxin axit nucleic này có thể ở dạng bất kỳ, bao gồm và trần, plasmid, hoặc các vector virus bao gồm các vector adenovirus.

Việc sử dụng polypeptit, các phân tử axit nucleic, và/hoặc các chế phẩm gây miễn dịch theo sáng chế có thể được thực hiện bằng các đường sử dụng tiêu chuẩn. Các ví dụ không giới hạn bao gồm sử dụng ngoài đường tiêu hóa, như trong mạch, trong da, qua da, trong cơ, dưới da, v.v., hoặc sử dụng ở niêm mạc, ví dụ trong mũi, miệng và tương tự. Người có hiểu biết trung bình trong lĩnh vực có thể xác định các khả năng khác nhau để sử dụng polypeptit, các phân tử axit nucleic, và/hoặc các chế phẩm gây miễn dịch theo sáng chế, để gây đáp ứng miễn dịch. Theo một số phương án, polypeptit, phân tử axit nucleic, và/hoặc chế phẩm gây miễn dịch (hoặc vacxin) được sử dụng nhiều hơn một lần, tức là trong phác đồ được gọi là khơi mào-thúc đẩy cùng loài. Theo một số phương án trong đó polypeptit, phân tử axit nucleic, và/hoặc chế phẩm gây miễn dịch được sử dụng nhiều hơn một lần, việc sử dụng liều thứ hai có thể được thực hiện sau một khoảng thời gian, ví dụ, một tuần hoặc lâu hơn sau khi sử dụng liều thứ nhất, hai tuần hoặc lâu hơn sau khi sử dụng liều thứ nhất, ba tuần hoặc lâu hơn sau khi sử dụng liều thứ nhất, một tháng hoặc lâu hơn sau khi sử dụng liều thứ nhất, sáu tuần hoặc lâu hơn sau khi sử dụng liều thứ nhất, hai tháng hoặc lâu hơn sau khi sử dụng liều thứ nhất, 3 tháng hoặc lâu hơn sau khi sử dụng liều thứ nhất, 4 tháng hoặc lâu hơn sau khi sử dụng liều thứ nhất, v.v., đến vài năm sau khi sử dụng liều thứ nhất của polypeptit, phân tử axit nucleic, và/hoặc chế phẩm gây miễn dịch. Cũng có thể sử dụng vacxin nhiều hơn hai lần, ví dụ ba lần, bốn lần, v.v., do đó việc sử dụng khơi mào đầu tiên được theo sau bởi nhiều hơn một lần sử dụng tăng cường. Theo các phương án khác, polypeptit, phân tử axit nucleic, và/hoặc chế phẩm gây miễn dịch theo sáng chế được sử dụng chỉ một lần.

Polypeptit, các phân tử axit nucleic, và/hoặc các chế phẩm gây miễn dịch cũng có thể được sử dụng, hoặc ở dạng khơi mào, hoặc ở dạng tăng cường, trong phác đồ khơi mào-tăng cường khác loài.

Sáng chế còn đề xuất các phương pháp phòng ngừa và/hoặc điều trị bệnh do virus cúm ở đối tượng sử dụng polypeptit, các axit nucleic và/hoặc các chế phẩm được mô tả ở đây. Theo phương án cụ thể, phương pháp phòng ngừa và/hoặc điều trị bệnh do virus cúm ở đối tượng bao gồm việc sử dụng cho đối tượng cần điều trị lượng có hiệu quả của polypeptit, axit nucleic và/hoặc chế phẩm gây miễn dịch, như được mô tả

trên đây. Lượng có hiệu quả điều trị chỉ lượng polypeptit, axit nucleic, và/hoặc chế phẩm như được xác định ở đây, mà có hiệu quả phòng ngừa, làm thuyên giảm và/hoặc điều trị bệnh hoặc tình trạng bệnh lý bắt nguồn từ việc nhiễm virus cúm A nhóm 1 hoặc 2, và/hoặc virus cúm B. Phòng ngừa bao gồm việc làm ức chế hoặc làm giảm sự phát tán của virus cúm hoặc ức chế hoặc làm giảm sự khởi phát, phát triển hoặc tiến triển của một hoặc nhiều triệu chứng đi kèm với nhiễm virus cúm. Việc làm thuyên giảm như được sử dụng ở đây có thể chỉ việc làm thuyên giảm các triệu chứng bệnh nhìn thấy được hoặc quan sát được, sự xâm nhiễm và phát triển trong máu của virus (viremia), hoặc sự biểu hiện khác có thể đánh giá được của nhiễm virus cúm.

Những người cần điều trị bao gồm những người đã bị nhiễm có tình trạng bệnh lý do nhiễm virus cúm A nhóm 1 hoặc nhóm 2, hoặc virus cúm B, cũng như những người đã được phòng ngừa lây nhiễm virus cúm. Polypeptit, các axit nucleic và/hoặc các chế phẩm theo sáng chế do đó có thể được sử dụng cho đối tượng chưa từng phơi nhiễm, tức là, đối tượng không mắc bệnh gây ra bởi nhiễm virus cúm hoặc chưa từng và hiện nay không bị nhiễm virus cúm, hoặc cho đối tượng hiện nay đã và/hoặc đã từng nhiễm virus cúm.

Theo một phương án, phòng ngừa và/hoặc điều trị có thể được nhằm vào nhóm người bệnh dễ bị nhiễm virus cúm. Các nhóm người bệnh này bao gồm, nhưng không chỉ giới hạn ở ví dụ, người già (ví dụ ≥ 50 tuổi, ≥ 60 tuổi, và tốt hơn ≥ 65 tuổi), người trẻ (ví dụ ≤ 5 tuổi, ≤ 1 tuổi), người bệnh nằm viện và người bệnh đã từng được điều trị bằng hợp chất kháng virus nhưng có biểu hiện đáp ứng kháng virus không đầy đủ.

Theo phương án khác, polypeptit, các axit nucleic và/hoặc các chế phẩm gây miễn dịch có thể được sử dụng cho đối tượng kết hợp với một hoặc nhiều tác nhân hoạt tính khác, như vắc xin cúm hiện có hoặc trong tương lai, kháng thể đơn dòng và/hoặc tác nhân kháng virus, và/hoặc tác nhân kháng khuẩn, và/hoặc tác nhân điều biến miễn dịch. Một hoặc nhiều tác nhân hoạt tính khác có thể có lợi trong điều trị và/hoặc phòng ngừa bệnh do virus cúm hoặc có thể làm thuyên giảm triệu chứng hoặc tình trạng bệnh lý đi kèm với bệnh do virus cúm. Theo một số phương án, một hoặc nhiều tác nhân hoạt tính khác là thuốc giảm đau, thuốc chống sốt, hoặc các thuốc điều trị làm giảm hoặc hỗ trợ thở.

Chế độ liều dùng của polypeptit và/hoặc các phân tử axit nucleic theo sáng chế có thể được điều chỉnh để mang lại đáp ứng tối ưu mong muốn (ví dụ, đáp ứng điều trị bệnh). Khoảng liều dùng thích hợp có thể chẳng hạn là từ 0,1-100 mg/kg khối lượng cơ thể, tốt hơn là từ 1-50 mg/kg khối lượng cơ thể, tốt hơn là từ 0,5-15 mg/kg khối lượng cơ thể. Liều dùng chính xác của polypeptit và/hoặc các phân tử axit nucleic để

được sử dụng sẽ ví dụ phụ thuộc vào đường sử dụng, và mức độ trầm trọng của tình trạng nhiễm hoặc bệnh gây ra bởi nhiễm virus, và phải được quyết định theo hướng dẫn của bác sĩ và tình trạng của đối tượng. Ví dụ, liệu có hiệu quả phụ thuộc nhiều vào vị trí đích, trạng thái sinh lý của người bệnh (bao gồm độ tuổi, khối lượng cơ thể, sức khỏe), và việc điều trị là dự phòng hoặc chữa bệnh. Thông thường, người bệnh là người nhưng cũng có thể là động vật có vú không phải người bao gồm động vật có vú chuyển gen cũng có thể được sử dụng. Liệu điều trị được chuẩn độ tối ưu để tối ưu hiệu lực và độ an toàn.

Polypeptit theo sáng chế cũng có thể được sử dụng để xác nhận liên kết của kháng thể đơn dòng được xác định là ứng viên có tiềm năng chữa bệnh. Ngoài ra, polypeptit theo sáng chế có thể được sử dụng làm công cụ chẩn đoán, ví dụ để kiểm tra tình trạng miễn dịch của cá thể bằng cách chứng minh liệu có kháng thể trong huyết thanh của các thể này có khả năng liên kết với polypeptit theo sáng chế. Do đó, sáng chế còn đề xuất phương pháp chẩn đoán in vitro để phát hiện sự có mặt của nhiễm virus cúm ở người bệnh, phương pháp này bao gồm các bước sau a) cho mẫu sinh học thu được từ người bệnh này tiếp xúc với polypeptit theo sáng chế; và b) phát hiện sự có mặt của phức hợp kháng thể-kháng nguyên.

Polypeptit theo sáng chế cũng có thể được sử dụng để nhận diện các phân tử liên kết mới hoặc cải thiện các phân tử liên kết hiện có, như kháng thể đơn dòng và tác nhân kháng virus.

Sáng chế được minh họa cụ thể hơn bằng các ví dụ và hình vẽ sau đây. Các ví dụ không nhằm làm giới hạn phạm vi bảo hộ của sáng chế theo cách bất kỳ.

Ví dụ thực hiện sáng chế

Ví dụ 1: Nhận diện kháng thể trung hòa chéo mới thuộc nhóm 1 và nhóm 2: CR9114

Máu ngoại vi được lấy từ người hiến máu khỏe mạnh bình thường bằng cách tiêm tĩnh mạch trong ống đựng mẫu chống đông chứa EDTA. Thư viện trung bày thực khuẩn thể scFv được thu nhận về cơ bản như được mô tả trong WO 2008/028946, được kết hợp bằng cách viện dẫn ở đây. Tiến hành chọn lọc kháng hemagglutinin (HA) tái tổ hợp của typ phụ H1 của virus cúm A (A/New Caledonia/20/99), H3 (A/Wisconsin/67/2005), H4 (A/Duck/Hong Kong/24/1976), H5 (A/Chicken/Vietnam/28/2003), H7 (A/Netherlands/219/2003) và H9 (A/HongKong/1073/99). Tiến hành hai vòng chọn lọc liên tiếp trước khi phân lập kháng thể chuỗi đơn của thực khuẩn thể tách biệt. Sau vòng chọn lọc thứ hai, các dòng *E.coli* tách biệt được sử dụng để tạo ra kháng thể đơn dòng của thực khuẩn thể.

Dịch nổi được chọn lọc chứa kháng thể chuỗi đơn của thực khuẩn thể mà thu được trong quá trình sàng lọc được mô tả trên đây được công nhận về tính đặc hiệu trong thử nghiệm ELISA, *tức là* liên kết với các kháng nguyên HA khác nhau. Để thực hiện mục đích này, baculovirut biểu hiện các HA tái tổ hợp của H1 (A/New Caledonia/20/99), H3 (A/Wisconsin/67/2005), H5 (A/Vietnam/1203/04) H7 (A/Netherlands/219/2003), và B (B/Ohio/01/2005) (Protein Sciences, CT, USA) được phủ Maxisorp™ lên đĩa ELISA. Trong số các kháng thể chuỗi đơn của thực khuẩn thể thu được, kháng thể chuỗi đơn của thực khuẩn thể SC09-114 thể hiện liên kết đặc hiệu với HA tái tổ hợp của virut cúm A typ phụ H1, H3, H5, H7 và virut cúm B. Liên kết và tính đặc hiệu của SC09-114 được xác nhận bằng phân tích FACS. Để thực hiện mục đích này, các HA tái tổ hợp chiều dài đủ của virut cúm A các typ phụ H1 (A/New Caledonia/20/1999), H3 (A/Wisconsin/67/2005) và H7 (A/Netherlands/219/2003) HA được biểu hiện trên bề mặt của tế bào PER.C6. SC09-114 được thể hiện liên kết đặc hiệu với HA của virut cúm A các typ phụ H1, H3 và H7. Các vùng biến đổi của chuỗi nặng và chuỗi nhẹ của scFv được tách dòng như được mô tả trước (WO 2008/028946). Các cấu trúc biểu hiện thu được mã hóa chuỗi nặng và nhẹ của IgG1 của người được biểu hiện tạm thời ở dạng kết hợp trong tế bào 293T và dịch nổi chứa kháng thể IgG1 của người CR9114 được thu nhận và tạo ra bằng cách sử dụng các quy trình tinh chế chuẩn. Trình tự axit amin của các CDR của chuỗi nặng và chuỗi nhẹ của CR9114 được đưa ra trong Bảng 1. Trình tự nucleotit và axit amin của các vùng biến đổi của chuỗi nặng và chuỗi nhẹ được đưa ra dưới đây.

Ví dụ 2: Hoạt tính liên kết chéo của CR9114

CR9114 được xác nhận về tính đặc hiệu liên kết trong ELISA, *tức là* liên kết với các kháng nguyên HA khác nhau. Để thực hiện mục đích này, baculovirut biểu hiện các HA tái tổ hợp của typ phụ H1 (A/New Caledonia/20/1999), H3 (A/Wisconsin/67/2005), H5 (A/Vietnam/1203/04), H7 (A/Netherlands/219/2003) và H9 (A/HongKong/1073/99) (Protein Sciences, CT, USA) được phủ Maxisorp™ lên đĩa ELISA. Để làm đối chứng, IgG không liên quan CR4098 được sử dụng. CR114 được thể hiện là có hoạt tính liên kết chéo khác typ phụ với tất cả các HA tái tổ hợp được thử nghiệm. Xem Bảng 2.

Ngoài ra, kháng thể được thử nghiệm về tính liên kết khác typ phụ bằng phân tích FACS. Để thực hiện mục đích này, các HA chiều dài đủ tái tổ hợp của virut cúm A các typ phụ H1 (A/New Caledonia/20/1999), H3 (A/Wisconsin/67/2005) và H7 (A/Netherlands/219/2003) được biểu hiện trên bề mặt của tế bào PER.C6. Ủ tế bào với CR9114 trong 1 giờ, sau đó là ba bước rửa bằng PBS+0,1%BSA. Kháng thể liên kết

được phát hiện nhờ sử dụng kháng thể liên hợp PE kháng người. Để làm đối chứng, tế bào PER.C6 không chuyển nhiễm được sử dụng. CR9114 thể hiện hoạt tính liên kết chéo với HA của virus cúm A các typ phụ H1, H3 và không liên kết với tế bào PER.C6 kiểu đại. Xem Bảng 2.

Ví dụ 3: Hoạt tính trung hòa chéo của CR9114

Để xác định xem CR9114 có khả năng phong bế nhiều chủng virus cúm A không, tiến hành các thử nghiệm trung hòa virus *in vitro* (VNA). Tiến hành VNA trên tế bào MDCK (ATCC CCL-34). Tế bào MDCK được nuôi cấy trong môi trường nuôi cấy tế bào MDCK (môi trường MEM được bổ sung kháng sinh, 20 mM Hepes và 0,15% (w/v) natri bicarbonat (môi trường MEM đầy đủ), được bổ sung 10% (v/v) huyết thanh bào thai bò). Các chủng thuộc typ phụ H1 (A/WSN/33, A/New Caledonia/20/1999, A/Solomon Islands/IVR-145 (chủng sắp xếp lại sinh trưởng cao của A/Solomon Islands/3/2006), A/Brisbane/59/2007, A/NYMC/X-181 (chủng sắp xếp lại sinh trưởng cao của A/California/07/2009), H2 (A/Env/MPU3156/05), H3 (A/Hong Kong/1/68, A/Johannesburg/33/94, A/Panama/2000/1999, A/Hiroshima/52/2005, A/Wisconsin/67/2005 và A/Brisbane/10/2007), H4 (A/WF/HK/MPA892/06), H5 (PR8-H5N1-HK97 (6:2 chủng sắp xếp lại của A/Hong Kong/156/97 và A/PR/8/34) và A/Eurasian Wigeon/MPF461/07), H6 (A/Eurasian Wigeon/MPD411/07), H7 (NIBRG-60 (6:2 chủng sắp xếp lại của A/Mallard/Netherlands/12/2000) và PR8-H7N7-NY (7:1 chủng sắp xếp lại của A/New York/107/2003 (H7N7) và A/PR/8/34)), H8 (A/Eurasian Wigeon/MPH571/08) H9 (A/Hong Kong/1073/99 và A/Chick/HK/SSP176/09), H10 (A/Chick/Germany/N/49) và H14 (PR8-H14N5 (6:2 chủng sắp xếp lại của A/mallard/Astrakhan/263/1982 (H14N5) và A/PR/8/34)) được sử dụng trong thử nghiệm tất cả được pha loãng đến độ chuẩn $5,7 \times 10^3$ TCID₅₀/ml (50% liều gây nhiễm nuôi cấy mô cho mỗi ml), với độ chuẩn được tính theo phương pháp của Spearman và Karber. Chế phẩm IgG (200 µg/ml) được pha loãng liên tiếp 2 lần (1:2 - 1:512) trong môi trường MEM đầy đủ lặp lại trong 4 giếng. 25 µl IgG mỗi độ pha loãng được trộn với 25 µl huyền phù chứa virus (100 TCID₅₀/25 µl) và được ủ trong 1 giờ ở nhiệt độ 37°C. Sau đó chuyển huyền phù 4 lần lên đĩa 96 giếng chứa dịch nuôi cấy MDCK nhập dòng trong 50 µl môi trường MEM đầy đủ. Trước khi sử dụng, tế bào MDCK được nuôi cấy ở mật độ 3×10^4 tế bào mỗi giếng trong môi trường nuôi cấy tế bào MDCK, được sinh trưởng đến khi tế bào đạt được nhập dòng, được rửa bằng 300-350 µl PBS, pH 7,4 và cuối cùng bổ sung 50 µl môi trường MEM đầy đủ vào mỗi giếng. Các tế bào được ủ được nuôi cấy

trong 3-4 ngày ở nhiệt độ 37°C và được quan sát hàng ngày về sự phát triển tác động gây bệnh tế bào (cytopathogenic effect-CPE). CPE được so với đối chứng dương.

CR9114 được thể hiện là có hoạt tính trung hòa chéo typ phụ khác ở các chủng đại diện trong tất cả các typ phụ virut cúm A được thử nghiệm H1, H2, H3, H4, H5, H6, H7, H8, H9 và H10 Xem Bảng 3.

Ví dụ 4: Thiết kế polypeptit vùng gốc chứa vùng gốc bảo thủ có các epitop của CR6261 và CR9114 dựa trên HA của H1

Kháng thể đơn dòng chiều dài đủ của người kháng hemagglutinin của virut cúm khả năng trung hòa chéo rộng đã được xác định trước đó. CR6261 (như được mô tả trong WO 2008/028946) được thể hiện là có hoạt tính trung hòa chéo rộng kháng virut cúm A thuộc nhóm phát sinh chủng loại 1. Ngoài ra, CR9114, được mô tả trên đây, đã được thể hiện là có khả năng liên kết với và trung hòa virut cúm A ở cả nhóm phát sinh chủng loại 1 và 2, cũng như virut cúm B. Phân tích về chức năng và cấu trúc đã chứng tỏ rằng các kháng thể này làm cản trở quá trình dung hợp màng và có hướng kháng epitop bảo thủ cao trong vùng gốc của protein HA của virut cúm (Throsby et al.(2008); Ekiert et al. (2009) WO2008/028946, và đơn yêu cầu cấp patent đang theo đuổi đồng thời số EP11173953.8).

Trong nghiên cứu mà dẫn đến sáng chế này, các phân tử mới chứa vùng gốc của HA chứa các epitop này được thiết kế để tạo ra các polypeptit gây miễn dịch dựa trên epitop phổ quát có thể được sử dụng, ví dụ làm vaccin gây ra sự bảo vệ kháng phổ rộng các chủng virut cúm. Quan trọng là, phần biến đổi cao và quyết định miễn dịch, tức là vùng đầu được loại bỏ đầu tiên từ phân tử HA chiều dài đủ để tạo ra polypeptit vùng gốc HA, còn được gọi là “mini-HA”. Theo cách này đáp ứng miễn dịch sẽ được chuyển hướng đến vùng gốc trong đó có các epitop cho các kháng thể trung hòa phổ rộng. Kháng thể CR6261 và CR9114 được sử dụng để thăm dò sự cuộn gấp chính xác của các phân tử mới được tạo ra, và để xác nhận sự có mặt của các epitop trung hòa.

Polypeptit theo sáng chế do đó trình diện các epitop bảo thủ của phân tử HA vùng cuống gần gốc màng cho hệ miễn dịch khi không có các epitop trội mà có mặt trong vùng đầu ở xa màng. Cuối cùng, một phần của trình tự ban đầu của protein HA0 tạo ra vùng đầu được loại bỏ và được nối lại, hoặc theo cách trực tiếp hoặc bằng cách đưa vào trình tự nối ngắn linh động(‘cầu nối’) để khôi phục tính liên tục của chuỗi. Trình tự thu được tiếp tục được cải biến bằng cách đưa các đột biến đặc hiệu làm ổn định cấu trúc ba chiều tự nhiên của phần còn lại của phân tử HA0.

Chức năng của phân tử HA trong virut là liên kết với axit sialic thụ thể bề mặt tế bào và, sau đó hấp thu vào trong endosom, làm trung gian cho sự dung hợp màng

virut và endosom dẫn đến sự giải phóng ARN của virut trong tế bào. Bước quan trọng trong quá trình dung hợp là sự thay đổi về cấu hình lớn của phân tử HA mà sắp xếp lại các yếu tố của cấu trúc bậc hai của phân tử này do đó peptit dung hợp trở nên được bộc lộ. Kết quả là, hai cấu hình (trước và sau dung hợp) của phân tử HA tồn tại rất khác nhau về cấu trúc bậc ba của chúng. Do protein HA của virut được bộc lộ chủ yếu cho hệ miễn dịch ở trạng thái trước dung hợp, điều quan trọng là đảm bảo rằng polypeptit theo sáng chế chấp nhận cấu hình này. Yêu cầu này có thể được đáp ứng bằng cách làm ổn định cấu hình trước dung hợp và đồng thời làm mất ổn định cấu hình sau khi dung hợp. Việc làm ổn định/mất ổn định này là cần thiết do cấu hình trước dung hợp là giả ổn định và chấp nhận cấu hình sau khi dung hợp tạo ra cấu hình phù hợp, tức là năng lượng tối thiểu (Chen et al, 1995).

Trong ví dụ này, HA từ H1N1 A/Brisbane/59/2007 (SEQ ID NO: 1) được lấy làm trình tự bậc 1 (hoặc kiểu đại) để tạo ra polypeptit theo sáng chế.

Trong bước thứ nhất, polypeptit theo sáng chế được tạo cấu trúc bằng cách loại bỏ các trình tự HA1 giữa các vị trí 59 và 291 (việc đánh số chỉ vị trí trong trình tự HA0, như được thể hiện trong SEQ ID NO: 1. Theo một số phương án, phần HA1 chứa axit amin 18-343 và phần HA2 chứa các gốc axit amin 344-565; do SEQ ID NO: 1 chứa peptit tín hiệu, và phần HA1 bắt đầu ở vị trí 18). Điều này dẫn đến việc loại bỏ các gốc 60 đến 290 của HA0. Các gốc này được thay thế bằng trình tự nối GGGG. Tiếp theo, diện tích bề mặt có thể truy nhập vào của mỗi gốc trong cả cấu hình trước và sau dung hợp được tính với sự hỗ trợ của Brugel (Delhaise et al., 1984). Mức độ bộc lộ và che giấu của mỗi gốc được xác định như được mô tả trong Samantha et al (2002), trong đó trọng tâm là các gốc mà được bộc lộ trong cấu hình trước dung hợp và bị che giấu trong cấu hình sau khi dung hợp. Các phân tích tiếp theo về các gốc này cho thấy rằng một số trong các gốc axit amin này có thể được gây đột biến theo cách mà đột biến này không có tác động lên cấu hình trước dung hợp nhưng làm mất ổn định cấu hình sau khi dung hợp. Các gốc này nhìn chung có chuỗi bên kỵ nước và tham gia vào sự tạo thành chuỗi cuộn xoắn trong cấu hình sau khi dung hợp. Việc gây đột biến các gốc axit amin thành axit amin ưa nước sẽ làm nhiều các đặc tính của chuỗi cuộn xoắn – các tương tác giữa các chuỗi xoắn ở dạng cuộn xoắn nhìn chung là tương tác kỵ nước – và do đó làm mất ổn định cấu hình sau khi dung hợp.

Sau lập luận này, trong phần HA2 của trình tự một số đột biến đã được đưa vào: Phe 406 thành Ser (F406S), Val 409 thành Thr (V409T), Leu 416 thành Ser (L416S) và Tyr 502 thành Ser (Y502S). Đây là các đột biến mà loại đi gốc kỵ nước từ bề mặt của HA. Cần lưu ý rằng đột biến L416 thành S hoặc T cũng đưa vào vị trí

glycosyl hóa N (trình tự liên ứng là NX(S/T)). Sự glycosyl hóa ở vị trí này sẽ làm tăng hơn nữa độ ổn định của vùng này. Ngoài ra, Leu 58 được đột biến thành Thr (L58T), Val 314 thành Thr (V314T) và Ile 316 thành Thr (I316T); các đột biến này đều nằm trong vùng HA1, tức là một phần của trình tự tương ứng với HA1 sau vị trí cắt của chuỗi HA0. Hai đột biến sau duy trì nhánh beta của chuỗi bên nhưng loại bỏ gốc kỵ nước từ bề mặt. Như được thấy dưới đây một số trong số các đột biến này được đưa vào trong tất cả các biến thể, các đột biến còn lại được thử nghiệm trong các polypeptit riêng rẽ để kiểm tra xem các đột biến có tác động lẫn nhau theo cách không mong muốn.

Để làm tăng độ ổn định của polypeptit, hai cầu disulfit được nghiên cứu để khóa HA trong cấu hình trước dung hợp. Cầu disulfit được tạo thành giữa các gốc ở khoảng cách thích hợp trong không gian với nhau (trong phân tử HA chiều dài đủ) và có nguyên tử C-beta của chúng ở vị trí chính xác để tạo ra cầu disulfit. Cầu disulfit đầu tiên được đề xuất là giữa vị trí 321 (vùng HA1) và vị trí 405 (vùng HA2). Nằm trong vùng HA2, cầu disulfit được tạo ra giữa các vị trí 413 và 421.

Do việc cắt phân tử HA ở vị trí R343 là bước cần thiết để sự thay đổi cấu hình có thể xảy ra, trong polypeptit theo sáng chế vị trí cắt này được loại bỏ bằng cách đưa vào đột biến Arg (R) thành Gln (Q). Một giải pháp khác theo sáng chế là thay đổi Arg thành Gln và loại bỏ các gốc 345 đến 350, một phần nhỏ của peptit dung hợp của HA2. Việc loại bỏ trình tự này (kỵ nước) sẽ làm ổn định hơn cho polypeptit.

Theo một số phương án, polypeptit theo sáng chế chứa trình tự nội bào của HA và vùng xuyên màng. Theo các phương án khác, trình tự trong tế bào chất và trình tự xuyên màng từ vị trí (hoặc dạng tương đương của nó) 523, 524, 525, 526, 526, 527, 528, 529, hoặc 530 của HA2 đến đầu tận cùng C của HA2 (đánh số theo trình tự SEQ ID NO: 1) được loại bỏ do đó polypeptit được tiết ra (hòa tan được) được tạo ra sau khi biểu hiện trong tế bào, có thể được sử dụng ví dụ trong vaccin. Polypeptit hòa tan được làm ổn định hơn nữa bằng cách đưa vào trình tự đã biết để tạo cấu trúc trime, tức là AYVRKDGEWVLL (SEQ ID NO: 143), tùy ý được nối thông qua cầu nối. Cầu nối này có thể tùy ý chứa vị trí phân cắt để xử lý sau đó theo quy trình mà người có hiểu biết trung bình trong lĩnh vực đã biết. Để dễ dàng tinh chế dạng hòa tan, có thể bổ sung thêm trình tự đuôi, ví dụ đuôi his (HHHHHHH) được nối thông qua cầu nối ngắn, ví dụ EGR. Theo một số phương án cầu nối và trình tự đuôi his được thêm vào không có trình tự cuộn gấp được thể hiện. Theo sáng chế, trình tự axit amin từ vị trí 530 (đánh số theo trình tự SEQ ID NO: 1) đến gốc axit amin đầu tận cùng C của vùng HA2 được loại bỏ và được thay thế bằng các trình tự sau:

EGRHHHHHHH (SEQ ID NO: 81), chứa cầu nối ngắn và đuôi his, hoặc SGRSLVPRGSPGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHH (SEQ ID NO: 82), chứa vị trí cắt thrombin, vùng trime hóa, và đuôi his.

Đột biến được mô tả trên đây được tạo nhóm thành cụm theo chức năng và vị trí của chúng trong cấu trúc ba chiều của các polypeptit vùng gốc HA. Tất cả các polypeptit chứa trình tự HA của H1 1-59 và 291-565 và đột biến R343Q, có thêm các đột biến sau đây: L58T, V314T, I316T, F406S, V409T, L416S (SEQ ID NO: 3; được gọi là cụm 1). Ngoài ra các biến thể cũng được tạo ra mà có thêm các đột biến sau:

Cụm 2: K321C, Q405C (SEQ ID NO: 4)

Cụm 3: F413C, E421C (SEQ ID NO: 5)

Cụm 4: HA2 Y502S (SEQ ID NO: 6)

Hơn nữa hai biến thể được tạo ra mà có trình tự cụm 1 và ngoài ra đột biến của cụm 2 và 3 (SEQ ID NO: 7) hoặc cụm 2, 3 và 4 (SEQ ID NO: 8).

Các gen mã hóa các trình tự protein nêu trên được tổng hợp và tách dòng vào vector biểu hiện pcADN2004 sử dụng các phương pháp thông thường mà người có hiểu biết trung bình trong lĩnh vực đã biết. Vì lý do so sánh trình tự chiều dài đủ (SEQ ID NO: 1) được bao gồm trong thử nghiệm cũng như trình tự được mô tả bằng Steel et al (2010) (H1-PR8-dH1; SEQ ID NO: 24), mà dựa trên trình tự H1N1 A/Puerto Rico/8/1934.

Huyền phù tế bào HEK293F (Invitrogen) (10^6 tế bào/ml, 30 ml) được chuyển nhiễm bằng vector biểu hiện (1 μ g/ml) sử dụng 40 μ l 293transfectin làm tác nhân chuyển nhiễm và để cho tiếp tục nhân lên trong 2 ngày. Tế bào được thu gom, được chia thành các phần phân ước (0,3 ml, khoảng 3×10^5 tế bào) và các phần phân ước được xử lý bằng huyết thanh đa dòng xuất hiện kháng HA của H1 để thăm dò sự biểu hiện hoặc kháng thể đơn dòng đặc hiệu HA (5 microgram/ml) và kháng thể thứ cấp được sử dụng để nhuộm. Sau đó tế bào được phân tích bằng phương pháp phân loại tế bào liên kết huỳnh quang (fluorescence associated cell sorting - FACS) về sự biểu hiện của polypeptit vùng gốc của HA được gắn lên màng theo sáng chế sử dụng huyết thanh đa dòng xuất hiện kháng HA của H1 để thăm dò sự biểu hiện. Dải kháng thể đơn dòng có tính đặc hiệu đã biết mà liên kết với protein chiều dài đủ (ví dụ CR6261 và CR9114) được sử dụng để thăm dò về sự có mặt của epitop bảo thủ và, bằng cách suy luận, sự cuộn gấp chính xác của HA chiều dài đủ và các mini-HA polypeptit theo sáng chế. Các kết quả được biểu hiện dưới dạng tỷ lệ phần trăm tế bào dương tính và cường độ phát huỳnh quang trung bình và được thể hiện trong FIG. 2.

Các kết quả thể hiện rằng tất cả các cấu trúc được biểu hiện trên bề mặt tế bào phản ứng với huyết thanh đa dòng (kháng-poly H1) dẫn đến 75% hoặc cao hơn trong số tất cả các tế bào được phân tích là dương tính so với trường hợp 4% đối với tế bào không được chuyển nhiễm. Kết quả này được xác nhận bằng các trị số của cường độ phát huỳnh quang trung bình (MFI), mà tương tự đối với tất cả các cấu trúc sau khi xử lý bằng huyết thanh đa dòng. Các thử nghiệm đối chứng không có IgG, chỉ sử dụng kháng IgG người được đánh dấu hoặc mAb không liên quan tất cả đều âm tính. Cả hai protein HA chiều dài đủ của chủng the A/Brisbane/59/2007 và A/Puerto Rico/8/1934 được nhận diện bằng kháng thể đơn dòng CR6261, CR6254, CR6328 (tất cả được biết là liên kết và trung hòa HA của H1; Throsby et al. (2008), WO 2008/028946), CR9114 (được mô tả trên đây), CR8001 (liên kết với HA của H1, nhưng không trung hòa H1; được mô tả trong WO2010/130636), nhưng CR8057 thì không (chỉ liên kết với một số chủng H3, cũng được mô tả trong WO2010/130636) và CR6307 (Throsby et al. (2008), WO 2008/028946).

Xem xét đặc điểm cấu hình và không liên tục của epitop CR6261 (Ekiert et al. 2009) kết luận rằng cả hai protein chiều dài đủ đều có mặt trong cấu hình 3 chiều tự nhiên của chúng. Đối với polypeptit được thiết kế mới theo sáng chế mà được thử nghiệm trong thử nghiệm này quan sát được thông số tương tự về sự nhận diện bởi dải kháng thể đơn dòng: liên kết bởi CR6261, CR6254, CR6328, CR9114 và CR8001 nhưng không liên kết bởi CR6307 và CR8057. Đây là bằng chứng rõ ràng nhất trong dữ liệu về tỷ lệ phần trăm tế bào dương tính, nhưng cũng quan sát được dữ liệu về cường độ phát huỳnh quang trung bình. Các kết quả tốt nhất thu được với miniHA-cụm1 cả về % tế bào dương tính cũng như cường độ phát huỳnh quang trung bình.

Những sự cải biến bổ sung hơn nữa, như cầu disulfit (cụm 2 và 3) được mô tả trên đây và đột biến Y502S của cụm 4 (hoặc sự kết hợp của các cải biến này, dẫn đến giảm tỷ lệ phần trăm của tế bào dương tính và giảm cường độ trung bình. Cấu trúc này được mô tả bởi Steel et al. (2010) (SEQ ID NO: 24) mà bao gồm việc loại bỏ vùng đầu, nhưng không có các cải biến hơn nữa không được nhận diện cao hơn mức độ nền bởi kháng thể bất kỳ được sử dụng trong thử nghiệm này. Do đó, kết luận rằng sau chuyển nhiễm bằng ADN protein này không được biểu hiện ở cấu hình 3 chiều tự nhiên mà có trong HA.

Các kết quả được mô tả trên đây chỉ về hướng tầm quan trọng của đột biến cụm 1 làm tăng đặc tính ưa nước của vòng được tạo bởi các gốc 402 đến 418 nối chuỗi cuộn xoắn A và chuỗi cuộn xoắn bộ khung dài (CD) của phân tử HA và vùng xung quanh. Để đưa vào hơn nữa tác động có lợi đột biến của cụm1 về độ ổn định và cuộn

gấp của polypeptit theo sáng chế miniHA (SEQ ID NO: 2; Polypeptit theo Steel, but dựa trên A/Brisbane) và miniHA_cụm1 (polypeptit theo sáng chế; SEQ ID NO: 3) được so sánh trong một thử nghiệm riêng (FIG. 3).

Trong khi khoảng 60% tế bào được chuyển nhiễm bằng miniHA-cụm1 là dương tính sau liên kết của CR6261, CR6254, CR6328 và CR9114, sự chuyển nhiễm bằng miniHA (polypeptit theo Steel, nhưng dựa trên A/Brisbane; SEQ ID NO: 2) dẫn đến các trị số rất gần với nồng độ nền (1-3%). Các tác giả sáng chế đã kết luận rằng đột biến của cụm1 góp phần có lợi vào sự cuộn gấp đúng cách và độ ổn định của polypeptit theo sáng chế, so với protein miniHA không được cải biến (SEQ ID NO: 2) mà không có các đột biến này.

Steel et al. tạo ra phân tử mới bằng cách loại bỏ gốc axit amin 53 đến 276 của HA1 của chủng H1 A/Puerto Rico/8/1934 và H3 HK68 ra khỏi trình tự bậc một, và thay thế đoạn bị loại bỏ này bằng cầu nối linh động. Như được thể hiện trong ví dụ này điều này dẫn đến phân tử không ổn định cao mà không điều chỉnh cấu hình chính xác, như được chứng tỏ do không có liên kết của các kháng thể mà được thể hiện trước đó là có liên kết với epitop bảo thủ trong vùng gốc. Sự cuộn gấp không đúng cách có nguyên nhân do sự tiếp xúc với dung môi của diện tích bề mặt lớn mà thường được che giấu bằng đầu hình cầu trong phân tử HA chiều dài đủ. Bởi vì vùng này có bản chất là kỵ nước, nên phân tử không còn ổn định và do đó sự thích nghi là cần thiết.

Sự chuyển đổi của các gốc kỵ nước thành các gốc ưa nước như được thực hiện trong polypeptit theo sáng chế kháng lại tác động này và làm ổn định polypeptit vùng gốc của HA. Đạt được sự làm ổn định hơn nữa cho sự cuộn gấp ba chiều tự nhiên của vùng gốc bằng cách đưa cầu disulfit vào vị trí thích hợp để nối chặt các gốc gần nhau trong không gian trong cấu trúc bậc ba tự nhiên nhưng tách rời trong cấu trúc bậc một.

Ví dụ 5: Khả năng sinh miễn dịch của polypeptit vùng gốc của HA trong Ví dụ

4

Để đánh giá khả năng sinh miễn dịch của polypeptit vùng gốc các con chuột được gây miễn dịch bằng vector biểu hiện mã hóa H1 chiều dài đủ từ A/Brisbane/59/2007 (SEQ ID NO: 1), miniHA-cụm1 (SEQ ID NO: 3), miniHA-cụm1+2 (SEQ ID NO: 4) và miniHA- cụm1+4 (SEQ ID NO:6). Vì lý do so sánh thiết kế miniHA bởi Steel et al. (2010) (mini-PR8; SEQ ID NO: 24) và protein HA chiều dài đủ tương ứng từ A/Puerto Rico/8/1934 cũng được bao gồm trong thử nghiệm này. Vector biểu hiện mã hóa cM2 cũng được bao gồm làm đối chứng âm.

Các nhóm gồm 4 con chuột (BALB/c) được gây miễn dịch bằng 50 µg cấu trúc + 50 µg tá được (pUMCV1-GM-CSF) tiêm bắp vào ngày 1, 21 và 42. Vào ngày 49

tiến hành lấy máu lần cuối và thu lấy huyết thanh. Phân tích huyết thanh bằng thử nghiệm FACS. Huyền phù tế bào HEK293F (Invitrogen) (10^6 tế bào/ml, 30 ml) được chuyển nhiễm bằng vector biểu hiện (1 microgram/ml) sử dụng 40 microlit 293transfectin làm tác nhân chuyển nhiễm và để cho tiếp tục nhân lên trong 2 ngày. Tế bào được thu gom, chia thành các phần phân ước (0,3 ml, khoảng 3×10^5 tế bào) và các phần phân ước được xử lý bằng huyết thanh đặc hiệu với cấu trúc này, được nhuộm bằng kháng thể thứ cấp và được phân tích bằng phương pháp phân loại tế bào liên kết huỳnh quang. Các kết quả được thể hiện trong FIG. 4.

Như mong đợi, huyết thanh đặc hiệu cM2 (đối chứng âm) nhận diện cM2, nhưng không nhận diện HA chiều dài đủ hoặc polypeptit vùng gốc như được chứng minh bằng % tế bào dương tính và MFI. Ngược lại, huyết thanh đặc hiệu HA chiều dài đủ nhuộm tế bào biểu hiện HA chiều dài đủ tương ứng (SEQ ID NO: 1), mà còn miniHA-cụm1 (SEQ ID NO: 3), miniHA-cụm1+2 (SEQ ID NO:4) và miniHA-cụm1+4 (SEQ ID NO:6), dù ở nồng độ thấp (ca 40% tế bào dương tính với ca 80% đối với chiều dài đủ, MFI ca 1000 với ca 7000 đối chiều dài đủ). Điều ngược lại cũng đúng: huyết thanh đặc hiệu đối với miniHA-cụm1 (SEQ ID NO: 3), miniHA-cụm1+2 (SEQ ID NO:4) và miniHA-cụm1+4 (SEQ ID NO:6) nhận diện tế bào biểu hiện cấu trúc tương ứng cũng như HA chiều dài đủ (SEQ ID NO: 1). Các kết quả được tóm tắt trong Bảng 4, dưới đây.

Ngược lại với kết quả trên đây đối với miniHA-cụm1 (SEQ ID NO: 3), miniHA-cụm1+2 (SEQ ID NO:4) và miniHA-cụm1+4 (SEQ ID NO:6), huyết thanh thu được từ chuột được gây miễn dịch bằng PR8 chiều dài đủ không liên kết tốt với tế bào được chuyển nhiễm bằng H1-PR8-dH1 (SEQ ID NO:24). Tỷ lệ phần trăm tế bào dương tính khoảng 20%, so với 40-50% đối với miniHA-cụm1 (SEQ ID NO: 3) và miniHA-cụm1+2 (SEQ ID NO: 4). Các kết quả còn được phản ánh trong cường độ phát huỳnh quang trung bình quan sát được mà chỉ vừa đủ cao hơn nồng độ nền.

Kết luận là, dữ liệu này cho thấy rằng polypeptit theo sáng chế có khả năng gây đáp ứng miễn dịch hướng đến HA chiều dài đủ. Cụ thể là các cải biến trong vùng giữa gốc 402 và 418 (đánh số theo trình tự SEQ ID NO: 1) là quan trọng để tạo ra phân tử bền vững.

Ví dụ 6: Tạo ra polypeptit vùng gốc thể hệ hai

Cường độ phát huỳnh quang trung bình đối với polypeptit vùng gốc được mô tả trong Ví dụ 4 trong tất cả các trường hợp đều thấp hơn quan sát được đối với protein chiều dài đủ tương ứng; thực tế thiết kế tốt nhất, miniHA-cụm 1 (SEQ ID NO: 3), có cường độ mà theo thứ tự là 10% cường độ trung bình của cấu trúc có chiều dài đủ sau

khi liên kết với kháng thể đơn dòng. Điều này chứng tỏ rằng sự biểu hiện và/hoặc cuộn gấp của polypeptit vùng gốc trên bề mặt tế bào là thấp hơn quan sát được đối với protein chiều dài đủ và rằng các thiết kế này có thể được cải thiện hơn nữa. Các kết quả thu được từ thể hệ thứ nhất thể hiện rằng sự cải thiện các cấu trúc thể hệ thứ nhất là có thể và do đó bắt đầu vòng thiết kế thứ hai.

Polypeptit được mô tả trong Ví dụ 4 dựa trên cùng đoạn bị mất của chuỗi HA0, tức là các gốc L60 đến K290 (mini1; đánh số chỉ vị trí trong HA0 chiều dài đủ từ H1N1 A/Brisbane/59/2007; SEQ ID NO: 1). Nghiên cứu này tạo ra vòng không có cấu trúc dài hiện tại không còn gắn với vùng đầu. Đó là lý do vì sao vòng này không góp phần vào độ ổn định của toàn bộ protein và có thể được làm ngắn đi một cách đáng kể mà không có tác động đến sự cuộn gấp của các phần khác của polypeptit. Ba đoạn được loại bỏ bổ sung được thiết kế và được thay thế bằng trình tự cầu nối GGGG như trước và có được kết hợp các đột biến của cụm1 được mô tả trên đây. Các đoạn được loại bỏ từ S53 đến P320 (mini2), H54 đến I302 (mini3), G56 đến G317 (mini4). Các cải biến bổ sung được đưa vào giống với cụm 1 trên đây (L58T, V314T, I316T, F406S, V409T, L416S).

Một số trong số các gốc thuộc cụm này là một phần của trình tự được loại bỏ và do đó có thể không được cải biến nữa (xem dưới đây). Hơn nữa, hai đột biến bổ sung được tạo ra trong chuỗi cuộn xoắn C dài mà tạo thành trime cuộn xoắn ở trạng thái trước dung hợp. Đã biết rõ trong lĩnh vực rằng các trime cuộn xoắn được làm ổn định bởi Ile ở các vị trí 420 a và d của trình tự lặp lại nhóm 7 mà là dấu chuẩn xác nhận của motif cấu trúc này (Suzuki et al., (2005); Woolfson et al. (2005)). Kiến thức này được áp dụng bằng cách đưa Ile vào các vị trí 420 (M420I) và 427 (V427I). Sự kết hợp của hai đột biến này và đột biến của cụm 1 được thiết kế cụm11; để làm rõ các sự kết hợp được đưa ra dưới đây:

Mini1: mất đoạn L60 đến K290 cụm11: M420I, V427I, L58T, V314T, I316T,

F406S, V409T, L416S

Mini2: mất đoạn S53 đến P320 cụm11: M420I, V427I, F406S, V409T, L416S

Mini3: mất đoạn H54 đến I302 cụm11: M420I, V427I, V314T, I316T, F406S,

V409T, L416S

Mini4: mất đoạn G56 đến G317 cụm11: M420I, V427I, F406S, V409T, L416S

Để làm ổn định hơn trạng thái trước dung hợp của polypeptit vùng gốc một cầu disulfit nữa được đưa vào giữa vị trí 324 và 436 (cụm 5: R324C, T436C) và được kết hợp với các thể đột biến mất đoạn khác. Các dạng kết hợp sau được tổng hợp và thử nghiệm về liên kết trong thử nghiệm FACS như được mô tả trên đây:

Mini1-cụm11 (SEQ ID NO: 9)

Mini2-cụm11 (SEQ ID NO: 10)

Mini3-cụm11 (SEQ ID NO: 11)

Mini4-cụm11 (SEQ ID NO: 12)

Mini1-cụm11+5 (SEQ ID NO: 13)

Mini2-cụm11+5 (SEQ ID NO: 14)

Mini3-cụm11+5 (SEQ ID NO: 15)

Mini4-cụm11+5 (SEQ ID NO: 16)

Vì lý do so sánh miniHA-cụm1 (SEQ ID NO:3) cũng được bao gồm trong thử nghiệm. Các kết quả được thể hiện trong FIG. 5.

Trong tất cả các trường hợp polypeptit vùng gốc có mặt trên bề mặt tế bào sau khi chuyển nhiễm vector biểu hiện vào tế bào HEK293F, như được chứng minh bằng tỷ lệ phần trăm của tế bào dương tính (90% hoặc lớn hơn) sau khi xử lý bằng huyết thanh kháng thể đa dòng kháng H1.

Tất cả polypeptit vùng gốc của HA trong thử nghiệm này được nhận diện bởi CR6261, CR6254, CR6328 và CR9114, nhưng không được nhận diện bởi CR8057; kháng thể sau là mong muốn bởi vì mAb này đặc hiệu đối với HA của H3. Tuy nhiên, có sự khác nhau rõ ràng về tỷ lệ phần trăm của tế bào dương tính và MFI đối với các kháng thể khác. Kháng thể được xác định đặc tính rõ nhất là CR6261, trong đó epitop này là đã biết chi tiết. Epitop này là gián đoạn và có cấu hình, và do đó, liên kết của kháng thể này có thể được coi là thử nghiệm nghiêm ngặt về sự cuộn gấp chính xác của polypeptit vùng gốc của HA. CR9114 trung hòa phổ rộng, bao gồm các chủng từ cả nhóm 1 và 2 (Bảng 3). Trong số các epitop của CR6328 và CR6254 chi tiết được thể hiện ít hơn, nhưng dựa trên các trị số cao hơn mà được tìm thấy đối với % tế bào dương tính và MFI, cũng như sự phát tán nhỏ hơn trong dữ liệu, liên kết của các kháng thể này dường là mẫu dò ít nhạy hơn của sự cuộn gấp chính xác so với CR6261.

So sánh tỷ lệ phần trăm tế bào dương tính (lấy trong báo cáo dữ liệu đối với toàn bộ kháng thể) các cấu trúc Mini1 đến 4 có thể được sắp xếp (cao nhất đến thấp nhất %).

Mini2 > Mini1 > Mini4 > Mini3 đối với các kết hợp với cụm11 và

Mini2 > Mini1 = Mini4 > Mini3 đối với các kết hợp với cụm11+5

Sự sắp xếp này có thể được bao gồm trong dữ liệu về MFI, và dẫn đến kết luận rằng việc loại bỏ cấu trúc Mini2, S53 thành P320, dẫn đến nồng độ protein cao nhất được biểu hiện trên bề mặt tế bào trong cấu hình chính xác từ tập hợp này.

So sánh MiniHA-cụm1 (SEQ ID NO: 3) với mini1-cụm11 (SEQ ID NO: 9), các đột biến bổ sung M420I, V427I dường như không dẫn đến việc làm ổn định hơn nữa cấu trúc này; nếu có, chúng dẫn đến tỷ lệ phần trăm tế bào dương tính và các trị số MFI thấp hơn, nhưng những sai khác này là nhỏ.

Việc đưa vào cầu disulfit R324C, T436C (cụm 5) dẫn đến làm tăng protein được cuộn gấp chính xác trên bề mặt tế bào đối với mini2-cụm11 (SEQ ID NO: 10) và mini4-cụm11 (SEQ ID NO: 12), nhưng cải thiện nhỏ hoặc không cải thiện đối với mini1-cụm11 (SEQ ID NO: 9) và mini3-cụm11 (SEQ ID NO: 11). Kết quả tổng thể tốt nhất thu được với mini2-cụm11+5 (SEQ ID NO: 14). Đây là bằng chứng cụ thể từ các trị số MFI mà đối với cấu trúc này là ca 50% của trị số đối với cấu trúc chiều dài đủ.

Theo một số phương án, polypeptit theo sáng chế chứa trình tự nội bào của HA và vùng xuyên màng. Theo các phương án khác, trình tự trong tế bào chất và trình tự xuyên màng từ vị trí (hoặc dạng tương đương của nó) 523, 524, 525, 526, 526, 527, 528, 529, hoặc 530 của HA2 đến đầu tận cùng C của HA2 (đánh số theo trình tự SEQ ID NO: 1) được loại bỏ, và tùy ý được thay thế bằng cách đưa vào trình tự đã biết để tạo cấu trúc trime, tức là AYVRKDGGEVLL (SEQ ID NO: 143), tùy ý được nối thông qua cầu nối. Cầu nối này có thể tùy ý chứa vị trí phân cắt để xử lý sau đó theo quy trình mà người có hiểu biết trung bình trong lĩnh vực đã biết. Để dễ dàng tinh chế dạng hạ tan có thể bổ sung thêm trình tự đuôi, ví dụ đuôi his HHHHHHHH được nối thông qua cầu nối ngắn, ví dụ EGR. Theo sáng chế, trình tự axit amin từ vị trí 530 (đánh số theo trình tự SEQ ID NO: 1) đến gốc axit amin đầu tận cùng C của vùng HA2 được loại bỏ và được thay thế bằng SEQ ID NO: 81 hoặc SEQ ID NO: 82.

Ví dụ 7: Khả năng sinh miễn dịch của polypeptit vùng gốc của HA thể hệ hai

Để đánh giá khả năng sinh miễn dịch của polypeptit vùng gốc trong Ví dụ 6, các con chuột được gây miễn dịch bằng vector biểu hiện mã hóa H1 chiều dài đủ từ A/Brisbane/59/2007 (SEQ ID NO: 1), miniHA-cụm1 (SEQ ID NO: 3), Mini2-cụm11 (SEQ ID NO: 10), Mini1-cụm11+5 (SEQ ID NO: 13), Mini2-cụm11+5 (SEQ ID NO: 14). Vector biểu hiện mã hóa cho cM2 cũng được bao gồm làm đối chứng âm.

Các nhóm gồm 4 con chuột (BALB/c) được gây miễn dịch bằng 50 µg cấu trúc + 50 µg tá dược (pUMCV1-GM-CSF) tiêm bắp vào ngày 1, 21 và 42. Vào ngày 49 tiến hành lấy máu lần cuối và thu lấy huyết thanh. HA0 chiều dài đủ (SEQ ID NO: 1), đối chứng âm cM2 và Mini2-cụm11+5 (SEQ ID NO: 14) cũng được sử dụng cho các

nhóm con chuột riêng rẽ bằng súng bắn gen, sử dụng ca 10 μ g cấu trúc + ca. 2 μ g tá được (pUMCV1-GM-CSF) và phác đồ gây miễn dịch tương tự. Huyết thanh được phân tích bằng ELISA sử dụng vùng ngoài của HA chiều dài đủ tái tổ hợp từ chủng A/Brisbane/59/2007 (lấy từ hãng Protein Sciences Corporation, Meriden, CT, USA) làm kháng nguyên. Tóm lại, phủ lên đĩa 96 giếng bằng 50 ng HA qua đêm ở 4 °C, sau đó ủ với đệm phong bế (100 μ l PBS, pH 7,4 + 2% sữa không kem) trong 1 giờ ở nhiệt độ trong phòng. Rửa các đĩa bằng PBS + 0,05% Tween-20, và 100 μ l của dãy độ pha loãng 2 lần trong đệm phong bế, bắt đầu từ độ pha loãng 20 lần của huyết thanh được bổ sung. Kháng thể liên kết được phát hiện bằng cách sử dụng IgG của dê kháng chuột liên hợp với HRP, sử dụng các quy trình chuẩn đã được thiết lập trong lĩnh vực này. Các độ chuẩn được so với đường cong chuẩn sử dụng mAb 3AH1 InA134 (Hytest, Turku, Finland) để nhận được số đơn vị/ml theo ELISA (EU/ml).

Các kết quả của ELISA sau 28 và 49 ngày được thể hiện trên Fig. 6A và B, tương ứng. Huyết thanh thu được từ các con chuột được gây miễn dịch bằng ADN mã hóa Mini2-cụm11+5 (SEQ ID NO: 14) biểu thị liên kết rõ ràng với vùng ngoài của HA chiều dài đủ sau 28 và 49 ngày sau khi gây miễn dịch sử dụng súng bắn gen và cũng sau 49 ngày sau khi gây miễn dịch trong cơ IM. For Mini2-cụm11 (SEQ ID NO: 10) và Mini1-cụm11+5 (SEQ ID NO: 13) phát hiện được đáp ứng ở 1 trong số 4 con chuột, trong khi đối với miniHA-cụm1 (SEQ ID NO: 3) không phát hiện được liên kết.

Kết luận là, dữ liệu này cho thấy rằng polypeptit theo sáng chế có khả năng gây đáp ứng miễn dịch hướng đến HA chiều dài đủ. Cụ thể là các cải biến trong vùng giữa gốc 402 và 418 (đánh số theo trình tự SEQ ID NO: 1), loại bỏ đoạn từ S53 đến P320 kết hợp với cầu disulfid R324C, T436C là quan trọng để tạo ra phân tử ổn định.

Ví dụ 8: Tạo ra polypeptit vùng gốc thể hệ ba

Để cải thiện hơn nữa thiết kế polypeptit vùng gốc vòng thiết kế thứ ba được thực hiện. Đột biến bổ sung để làm tăng tính ưa nước của các bề mặt được che giấu trong HA chiều dài đủ, nhưng không phải polypeptit vùng gốc được đưa vào ở vị trí 413, F413G (đánh số theo trình tự SEQ ID NO: 1), và được gọi là cụm 6. Cụm này được kết hợp với cụm được loại bỏ mini-2 (S53 đến P320), cầu disulfid của cụm 5 (R324C, T436C) và đột biến của cụm 1 (tức là F406S, V409T, L416S; SEQ ID NO: 46) hoặc cụm 11 (M420I, V427I, F406S, V409T, L416S; SEQ ID NO: 47). Kết hợp của việc loại bỏ đoạn mini-2 (S53 đến P320) với cụm 1 (F406S, V409T, L416S) và cụm 5 (R324C, T436C) cũng được bao gồm trong thử nghiệm này để tham khảo (SEQ ID NO: 48).

HA tự nhiên tồn tại ở dạng trime trên bề mặt tế bào. Hầu hết các tương tác giữa từng monome mà giữ dạng trime cùng nằm trong vùng đầu. Sau khi loại bỏ vùng đầu, do đó cấu trúc bậc ba được làm mất ổn định và vì vậy củng cố các tương tác giữa các monome trong phân tử bị cắt đứt sẽ làm tăng độ ổn định. Trong vùng gốc, trime hóa được trung gian bởi sự hình thành của môtip cuộn xoắn dạng trime. Bằng cách củng cố môtip này có thể tạo ra dạng trime thích hợp hơn. Theo sáng chế, trình tự liên ứng để tạo ra dạng trime xoắn cuộn, IEAIEKKIEAIEKKIE (SEQ ID NO: 83), được đưa vào trong polypeptit theo sáng chế ở (tương đương của) vị trí 418 đến 433 (SEQ ID NO: 44) trong chủng H1 A/Brisbane/59/2007 (đánh số theo trình tự SEQ ID NO: 1). Một cách khác là đưa vào trình tự MKQIEDKIEEIESKQ (SEQ ID NO: 84), có nguồn gốc từ GCN4 và được biết là trime hóa, ở vị trí 419-433 (SEQ ID NO: 45).

Trong trường hợp polypeptit vùng gốc được mô tả bằng trình tự SEQ ID NO: 44 với trình tự SEQ ID NO: 48 tất cả protein có mặt trên bề mặt tế bào sau khi chuyển nhiễm vector biểu hiện vào tế bào HEK293F, như được chứng minh bằng tỷ lệ phần trăm của tế bào dương tính (90% hoặc lớn hơn) sau khi xử lý bằng huyết thanh kháng thể đa dòng kháng H1. Các kết quả được thể hiện trên Fig.7.

Tất cả polypeptit vùng gốc của HA trong thử nghiệm này, ngoại trừ miniHA (SEQ ID NO: 2), được nhận diện bởi CR6261, CR6328 và CR9114, nhưng không được nhận diện bởi CR8020; kháng thể sau là mong muốn bởi vì mAb này đặc hiệu đối với HA của H3. Tỷ lệ phần trăm tế bào dương tính là khoảng 80% đối với polypeptit vùng gốc sử dụng CR6261, CR6328 và CR9114 để nhuộm, ngoại trừ miniHA, chỉ được nhận diện bởi huyết thanh kháng thể đa dòng kháng H1. Một lần nữa, điều này chứng tỏ rằng không có sự cuộn gấp đúng cách của cấu trúc cụ thể này. Tuy nhiên có sự sai khác rõ ràng về MFI đối với các kháng thể khác. Kháng thể được xác định đặc tính rõ nhất là CR6261, trong đó epitop này là đã biết chi tiết. Epitop này là gián đoạn và có cấu hình, và do đó, liên kết của kháng thể này có thể được coi là thử nghiệm nghiêm ngặt về sự cuộn gấp chính xác của polypeptit vùng gốc của HA. CR9114 trung hòa phổ rộng, bao gồm các chủng từ cả nhóm 1 và 2 (Bảng 3). Chi tiết của epitop của CR6328 được thể hiện ít hơn nhưng quan sát được trong các thử nghiệm liên kết về tính cạnh tranh của HA chiều dài đủ với CR6261.

MFI đối với H1-mini2-cl11+5 (SEQ ID NO: 14), H1-mini2-cl1+5 (SEQ ID NO: 48), H1-mini2-cl1+5+6 (SEQ ID NO: 46) và H1-mini2-cl11+5+6 (SEQ ID NO: 47) là rất giống nhau, bất chấp kháng thể đơn dòng được sử dụng trong thử nghiệm. Việc bao gồm vùng trime hóa liên ứng (SEQ ID NO: 44) làm giảm MFI bởi yếu tố 3 đến 4 so với trình tự tương đương không có vùng trime hóa (tức là H1-mini2-

cụm1+5+6; SEQ ID 46), nhưng kết quả vẫn là tốt hơn rõ ràng so với không có các cải biến đối với polypeptit gốc sau khi loại bỏ vùng đầu (các kết quả cf miniHA). Việc bổ sung trình tự trime hóa GCN4 (SEQ ID NO: 45) làm tăng nồng độ MFI có thể so sánh được với protein chiều dài đủ.

Ví dụ 9: Thiết kế polypeptit vùng gốc khác chứa vùng gốc bảo thủ có các epitop của CR6261 và CR9114

Polypeptit theo sáng chế được thiết kế theo quy trình được mô tả trên đây có thể được cải biến hơn nữa để làm tăng độ ổn định. Các cải biến này có thể được đưa vào để tăng cường sự hình thành các dạng trime của polypeptit theo sáng chế so với các loại monome và/hoặc đime. Như được mô tả trên đây, HA tự nhiên tồn tại ở dạng trime trên bề mặt tế bào. Hầu hết các tương tác giữa từng monome mà giữ dạng trime cùng nằm trong vùng đầu. Sau khi loại bỏ vùng đầu, do đó cấu trúc bậc ba được làm mất ổn định và vì vậy củng cố các tương tác giữa các monome trong phân tử bị cắt đứt sẽ làm tăng độ ổn định. Trong vùng gốc, trime hóa được trung gian bởi sự hình thành của môtip cuộn xoắn dạng trime. Bằng cách củng cố môtip này có thể tạo ra dạng trime thích hợp hơn.

Theo sáng chế, trình tự liên ứng để tạo ra dạng trime xoắn cuộn, IEAIEKKIEAIEKKIE (SEQ ID NO: 83), được đưa vào trong polypeptit theo sáng chế ở (tương đương của) vị trí 418 đến 433 (SEQ ID NO: 44) trong chủng H1 A/Brisbane/59/2007 (đánh số theo trình tự SEQ ID NO: 1). Theo cách khác IEAIEKKIEAIEKKI (SEQ ID NO: 85) có thể được đưa vào ở 419-433 (SEQ ID NO: 49) hoặc IEAIEKKIEAIEKK (SEQ ID NO: 86) ở 420-433 (SEQ ID NO: 50). Một cách khác là đưa vào trình tự MKQIEDKIEEIESKQ (SEQ ID NO: 84) có nguồn gốc từ GCN4 và được biết là trime hóa, ở vị trí 419-433 (SEQ ID NO: 45). Theo cách khác MKQIEDKIEEIESK (SEQ ID NO: 87) có thể được đưa vào ở vị trí 420-433 (SEQ ID NO: 51) hoặc RMKQIEDKIEEIESKQK (SEQ ID NO: 88) ở vị trí 417-433 (SEQ ID NO: 52). Tương tự, bề mặt trung gian của trime được củng cố bằng cách cải biến M420, L423, V427, G430 thành isoleuxin. (SEQ ID NO: 53).

Tất cả các peptit được thể hiện là liên kết với CR9114 và CR6261.

Theo một số phương án, polypeptit theo sáng chế không chứa trình tự tín hiệu và/hoặc các trình tự nội bào và vùng xuyên màng của HA, như được mô tả trước.

Ví dụ 10: Thiết kế polypeptit vùng gốc chứa vùng gốc bảo thủ có các epitop của CR6261 và CR9114 dựa trên H7 HA

Quy trình được mô tả trên đây để thiết kế polypeptit theo sáng chế cũng được áp dụng đối với H7. Trong ví dụ này việc thiết kế polypeptit theo sáng chế dựa trên

typ huyết thanh H7 được mô tả. HA của virus cúm H7 A/Mallard/Netherlands/12/2000 (SEQ ID NO: 31) được sử dụng làm trình tự bố mẹ, nhưng người có hiểu biết trung bình trong lĩnh vực sẽ hiểu rằng việc sử dụng các trình tự H7 khác là có thể tương đương bởi vì các trình tự có tính bảo thủ tốt, cụ thể là trong vùng gốc.

Cải biến thứ nhất trong trình tự là việc loại bỏ vị trí cắt ở vị trí 339 (đánh số đề cập đến trình tự SEQ ID NO: 31 bằng cách gây đột biến R thành Q (R339Q) để ngăn chặn sự tạo thành HA1 và HA2 từ HA0. Một cách tùy ý các gốc 341 đến 345 (LFGAI, một phần của peptit dung hợp) có thể được loại bỏ thêm để làm giảm đến mức thấp nhất sự tiếp xúc của các gốc kỵ nước với dung môi chứa nước. Điện tích dương ở vị trí cắt được bảo tồn 100% trong H7 và do đó đột biến này có thể được áp dụng trong tất cả các trình tự.

Cải biến thứ hai là việc loại bỏ vùng đầu bằng cách loại bỏ một phần lớn của trình tự HA1 và nối lại các trình tự đầu tận cùng N và C thông qua một cầu nối ngắn. Đoạn bị loại bỏ có thể thay đổi về chiều dài, nhưng được ưu tiên là gốc cuối cùng của trình tự đầu tận cùng N của HA1 và gốc đầu tiên của trình tự đầu tận cùng C gần nhau về mặt không gian để tránh việc đưa chúng vào thông qua trình tự nối. Trong các trình tự H7 các đoạn được loại bỏ có thể được đưa vào ở (các vị trí tương đương của) R53-P315 (mini2; SEQ ID NO: 33) trong chủng H7 A/Mallard/Netherlands/12/2000 (SEQ ID NO: 31). Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp ví dụ như Clustal hoặc Muscle. Các phần còn lại của trình tự có thể được nối trực tiếp hoặc theo cách khác có thể đưa vào cầu nối linh động. Trình tự cầu nối có thể có chiều dài từ 1 đến 50 axit amin. Được ưu tiên là các cầu nối linh động có chiều dài giới hạn (nhỏ hơn hoặc bằng 10 axit amin), ví dụ GGG, GGGG, GSA, GSAG, GSAGSA, GSAGSAG hoặc tương tự.

SEQ ID NO: 40 mô tả polypeptit này theo sáng chế chứa đoạn được loại bỏ từ T54-C314 (mini5; SEQ ID NO: 40). Các đoạn được loại bỏ được mô tả trên đây đảm bảo rằng các vùng không có cấu trúc được tạo thành bởi các gốc 280 – 310 cũng được loại bỏ; điều này có lợi đối với tính ổn định tổng thể của polypeptit theo sáng chế. Quan sát được tác động tương tự đối với polypeptit theo sáng chế có nguồn gốc từ trình tự H1 (xem trên đây).

Việc loại bỏ vùng đầu ra khỏi vòng giữa các gốc 394 đến 414 hiện được cho tiếp xúc với dung môi chứa nước. Trong H7, vòng của HA này có tính bảo thủ cao (xem Bảng 7). Trình tự liên ứng là: LI (E/D/G) KTNQQFELIDNEF (N/T/S) E (I/V) E (Q/K) (SEQ ID NO: 32).

Để làm tăng độ ổn định của vòng này trong cấu hình trước dung hợp và làm mất ổn định cấu hình sau khi dung hợp một số gốc kỵ nước được cải biến thành axit amin phân cực (S,T,N,Q), axit amin tích điện (R,H,K,D,E), hoặc làm gia tăng độ linh động bằng cách gây đột biến thành G. Gây đột đặc hiệu ở các vị trí 402, 404, 405, 409, 412 (đánh số đề cập đến trình tự SEQ ID NO: 31) sẽ góp phần làm ổn định polypeptit theo sáng chế.

Đối với các vị trí F402 và F409 đột biến thành S được ưu tiên nhưng các axit amin phân cực (T,N,Q), axit amin tích điện (R,H,K,D,E) và axit amin linh động cao (G) khác sẽ có tác động tương tự. Đối với vị trí 404 (96% L), đột biến thành N hoặc S được ưu tiên; axit amin sau cũng có thể xảy ra trong tự nhiên, mặc dù ở tần suất thấp, và đột biến của vị trí này trong các trường hợp này không cần thiết. Axit amin phân cực (T, Q), axit amin tích điện (R,H,K,D,E) và axit amin linh động cao (G) khác sẽ có tác động tương tự. Đối với vị trí 405 (99% I) đột biến thành T hoặc D được ưu tiên. D cũng có trong tự nhiên và đột biến của vị trí này sau đó là không cần thiết. Axit amin phân cực (S, N, Q), axit amin tích điện (R, H, K,) và axit amin linh động cao (G) khác sẽ có tác động tương tự. Đối với vị trí 412 (I hoặc V) đột biến thành N được ưu tiên nhưng axit amin phân cực (S, T, Q), axit amin tích điện (R, H, K, D, E) hoặc axit amin linh động (G) khác là cũng có thể được. Do đó các polypeptit chứa ít nhất một đột biến được mô tả trên đây. Các sự kết hợp của nhiều hơn một đột biến cũng được áp dụng, như được thể hiện trong các trình tự SEQ ID NO: 34-39 và 41-43.

Để làm ổn định cấu hình trước dung hợp của polypeptit theo sáng chế liên kết cộng hóa trị giữa hai phần xa nhau trong các trình tự bậc một nhưng gần nhau trong cấu hình được cuộn gấp trước dung hợp được đưa vào. Cuối cùng cầu disulfit được thiết kế trong polypeptit theo sáng chế, tốt hơn là ở giữa (tương đương của) vị trí 319 và 432 trong chủng H7A/Mallard/Netherlands/12/2000 (SEQ ID NO: 36-39, 42, 43). Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp như Clustal, Muscle etc. Cầu disulfit được thiết kế được tạo ra bằng cách gây đột biến ít nhất một (nếu gốc còn lại là xystein rồi), nhưng thường là hai gốc gần nhau về không gian với xystein, mà sẽ tạo thành liên kết cộng hóa trị một cách tự phát hoặc bằng phản ứng oxi hóa hoạt hóa giữa các nguyên tử lưu huỳnh của các gốc này.

Như được mô tả trên đây, HA tự nhiên tồn tại ở dạng trime trên bề mặt tế bào. Hầu hết các tương tác giữa từng monome mà giữ dạng trime cùng nằm trong vùng đầu. Sau khi loại bỏ vùng đầu, do đó cấu trúc bậc ba được làm mất ổn định và vì vậy củng cố các tương tác giữa các monome trong phân tử bị cắt cụt sẽ làm tăng độ ổn

định. Trong vùng gốc, trime hóa được trung gian bởi sự hình thành của môtip cuộn xoắn dạng trime. Bằng cách củng cố môtip này có thể tạo ra dạng trime thích hợp hơn. Đã biết rõ trong lĩnh vực rằng các trime cuộn xoắn được làm ổn định bởi Ile ở các vị trí a và d của trình tự lặp lại nhóm 7 đó là dấu chuẩn xác nhận của môtip cấu trúc này. Trong bản mô tả, kiến thức này được áp dụng bằng cách đưa Ile vào (tương đương của) các vị trí 419, 423, 426 và 430 (SEQ ID NO: 38, 43). Theo cách khác trình tự liên ứng để tạo ra dạng trime xoắn cuộn, EAIEKKIEAI, được đưa vào ở (tương đương của) vị trí 417 đến 426 (SEQ ID NO: 39).

Các trình tự này (SEQ ID NO: 33-43) được đưa vào thử nghiệm phân loại tế bào liên kết huỳnh quang được mô tả trên đây. Tuy nhiên, không có liên kết của kháng thể đơn dòng CR8020, CR8043, CR9114 hoặc CR8957 có thể được phát hiện. Các tác giả sáng chế đã kết luận rằng các trình tự này không trình diện các epitop của các kháng thể này và do đó các protein như được trình diện trên màng tế bào không được cuộn gấp thành cấu trúc 3 chiều tự nhiên của nó.

Ví dụ 11: Thiết kế polypeptit vùng gốc chứa vùng gốc bảo thủ có các epitop của CR8020, CR8043 và CR9114 dựa trên HA của H3

Trong bước thứ nhất, trình tự đại diện polypeptit theo sáng chế được tạo cấu trúc theo cách tương tự như được mô tả bởi Steel và các đồng tác giả (Steel et al 2010) sử dụng HA từ H3 A/Wisconsin/67/2005 như trình tự bố mẹ (SEQ ID NO: 89). Đầu của HA được loại bỏ bằng cách loại bỏ một phần của HA1 từ axit amin D69 đến axit amin K292.

Các gốc này có thể được thay thế bằng 3 hoặc 4 Gly. Cầu nối 4 Gly được thử nghiệm bởi Steel và các đồng tác giả và đã đưa ra các kết quả tốt về sự biểu hiện và được chấp nhận ở đây để tạo ra mini-H3 (SEQ ID NO: 90). Để ngăn chặn sự cắt chuỗi polypeptit, bước xử lý sau dịch mã thông thường đối với protein HA chiều dài đủ, vị trí cắt ở vị trí 345 (arginin) được gây đột biến thành glutamin (R345Q).

Tiếp theo, diện tích bề mặt có thể truy nhập được của mỗi gốc trong cả mini-HA được tạo cấu trúc và cấu hình sau khi dung hợp được tính với sự hỗ trợ của Brugel. Mức độ bộc lộ và che giấu của mỗi gốc được xác định như được mô tả trong Samantha và các đồng tác giả (Samantha et al., 2002). Tập chung vào các gốc mà được để lộ trong cấu hình trước dung hợp và trở nên được che giấu trong cấu hình sau khi dung hợp. Tiếp tục phân tích các gốc này đã chỉ ra rằng một số trong số các gốc này có thể được cải biến theo cách mà đột biến không có tác động đến cấu hình trước dung hợp nhưng có làm mất ổn định cấu hình sau khi dung hợp. Nhìn chung các gốc này có chuỗi bên kỵ nước và tham gia vào sự tạo thành chuỗi cuộn xoắn trong cấu

hình sau khi dung hợp. Đột biến của các gốc này để bao gồm chuỗi bên ưa nước sẽ phá hủy các đặc tính của chuỗi cuộn xoắn – tương tác giữa các cấu trúc xoắn trong chuỗi cuộn xoắn nhìn chung là tương tác kỵ nước – và do đó làm mất ổn định cấu hình sau khi dung hợp. Các gốc mà chuyển đổi từ trạng thái được bộc lộ trong cấu hình trước dung hợp thành được che giấu trong cấu hình sau khi dung hợp và được mong muốn rằng có tác động làm mất ổn định trên cấu hình sau, sau đột biến ở vị trí L397, I401 và L425 (đánh số theo trình tự SEQ ID NO: 89). Trong sáng chế L397K và I401T được bao gồm.

Vòng (vòng B, các gốc 401 đến 420) nối chuỗi cuộn xoắn A (gốc 383 đến 400) với chuỗi cuộn xoắn CD ở tâm (các gốc 421 đến 470) làm thay đổi cấu hình khi chấp nhận tình trạng sau dung hợp; nó trở thành vòng xoắn và là một phần của chuỗi cuộn xoắn trime kéo dài. Để làm ổn định cấu hình vòng trước dung hợp chứa cấu nối này và/hoặc để làm mất ổn định cấu hình sau dung hợp của nó được lý giải rằng nó cần phải đủ để gây đột biến tất cả các gốc mà tham gia vào việc tạo lõi của vòng cuộn xoắn. Đối với vị trí N405 một số đột biến được thiết kế, cụ thể là các gốc mang điện tích âm (Asp và Glu, N405D, N405E) bởi vì điện tích ngoài này sẽ củng cố nehaik ion quan sát thấy trong cấu hình trước khi dung hợp. Đột biến thành Ala trung hòa (N405A) cũng được bao gồm trong nghiên cứu này. Các tác giả sáng chế cũng gây đột biến Phe 408 thành Thr, His 409 thành Ser và Val 418 thành Ser (đánh số theo trình tự SEQ ID NO: 89; F408T, H409S, V418S) để làm tăng hơn nữa độ ổn định của bề mặt được bộc lộ mới sau khi loại bỏ vùng đầu.

Năm cầu disulfit được thiết kế để khóa HA trong cấu hình trước dung hợp. Các cầu này được tạo thành giữa các gốc ở các vị trí xa nhau thích hợp về không gian và đã có các nguyên tử C β của chúng ở vị trí chính xác để tạo thành cầu disulfit. Các cầu này được đưa vào giữa các vị trí 320 và 406 (A320C, E406C; đánh số theo trình tự SEQ ID NO: 89), 326 và 438 (K326C, S438C) và giữa 415 và 423 (F415C, Q423C). Hai cầu đầu tiên là liên kết chéo giữa các phần HA1 và HA2 của chuỗi, trong khi cầu cuối liên kết cộng hóa trị với phần trên cùng của vòng B. Cầu disulfit K326C, S438C được thực hiện bằng đột biến của Asp 435 thành Ala (D435A). Cầu disulfit F347C/N461C và S385C/L463C được lấy từ bài báo của Bommakanti và các đồng tác giả (2010) và cũng được sử dụng trong nghiên cứu này.

Để loại bỏ các gốc kỵ nước được bộc lộ mới từ dung môi, một vài đột biến bổ sung được thiết kế. Ile ở vị trí 67 (đánh số theo trình tự SEQ ID NO: 89) được gây đột biến thành Thr (I67T). Đột biến này duy trì nhánh beta của chuỗi bên nhưng loại bỏ các gốc kỵ nước ra khỏi bề mặt. Điều tương tự cũng có thể đối với đột biến Ile 298

thành Thr (I298T). Đột biến khác được đưa vào ở vị trí 316, isoleuxin trong trình tự tự nhiên. Bằng trực giác, một người có thể đề xuất tạo đột biến gốc này thành Thr để duy trì nhánh beta nhưng loại bỏ tính kỵ nước khỏi bề mặt. Tuy nhiên, đột biến này có thể dẫn đến việc đưa vào vị trí N-glycosyl hóa bổ sung (vị trí 314 là Asn) và do đó đột biến thành Gln được đưa vào (I316Q).

Gly 495 cũng được gây đột biến thành Glu (G495E). Đột biến này được thiết kế để đưa vào cầu ion bởi vì có điện tích dương trong vùng xung quanh. Nature đã đề xuất một số chủng H3 có Glu ở vị trí này.

Một gốc quan trọng của HA là vị trí 345 (Arg) bởi vì đây là vị trí trong đó xảy ra sự cắt bằng protaza để làm cho sự dung hợp protein có đủ khả năng. Đột biến của Arg này thành Gln (R345Q) ngăn chặn xảy ra việc phân cắt do đó khóa protein trong trạng thái trước dung hợp.

Đột biến được mô tả trên đây được tập hợp như được mô tả dưới đây:

Cụm 1:I67T, I98T, I316Q, F408T, H409S, V418S

Cụm 2: A320C, E406C

Cụm 3:K326C, D435A, S438C

Cụm 4:L397K, I401T

Cụm 5:N405D hoặc N405E hoặc N405A

Cụm 6:F415C, Q423C

Cụm 7:G495E

Cụm 8:F347C, S385C, N461C, L463C

Để đi đến polypeptit theo sáng chế, các cụm được kết hợp với việc loại bỏ đoạn D69 đến K292 và đột biến R345Q theo sơ đồ được mô tả dưới đây

H3 Mini-HA cụm 1 (SEQ ID NO: 91)

H3 Mini-HA cụm 1+2 (SEQ ID NO: 92)

H3 Mini-HA cụm 1+3 (SEQ ID NO: 93)

H3 Mini-HA cụm 1+4 (SEQ ID NO: 94)

H3 Mini-HA cụm 1+5 N405A (SEQ ID NO: 95)

H3 Mini-HA cụm 1+5 N405D (SEQ ID NO: 96)

H3 Mini-HA cụm 1+5 N405E (SEQ ID NO: 97)

H3 Mini-HA cụm 1+6 (SEQ ID NO: 98)

H3 Mini-HA cụm 1+7 (SEQ ID NO: 99)

H3 Mini-HA cụm 1+2+3+4+5+6+7-N405E (SEQ ID NO: 100)

H3 Mini-HA cụm 1+2+3+4+5+6+7-N405A (SEQ ID NO: 101) H3 Mini-HA cụm 1+2+3+4+5+6+7-N405D (SEQ ID NO: 102)

H3 Mini-HA cụm 1+8 (SEQ ID NO: 103).

Các gen mã hóa trình tự protein nêu trên được tổng hợp và tách dòng vào vector biểu hiện pcADN2004 sử dụng các phương pháp thông thường mà người có hiểu biết trung bình trong lĩnh vực đã biết. Vì lý do so sánh trình tự HA chiều dài đủ của H3 A/Wisconsin/67/2005 được bao gồm trong thử nghiệm, cũng như trình tự HA chiều dài đủ của H1 A/Brisbane/59/2007 chứa đột biến vị trí cắt R343Q.

Huyền phù tế bào HEK293F (Invitrogen) (10^6 tế bào /ml, 30 ml) được chuyển nhiễm bằng vector biểu hiện (1 μ g/ml) sử dụng 40 μ l 293transfectin làm tác nhân chuyển nhiễm và để cho tiếp tục nhân lên trong 2 ngày. Tế bào được thu gom, chia thành các phần phân ước (0,3 ml, khoảng 3×10^5 tế bào) và các phần phân ước được xử lý bằng huyết thanh đa dòng xuất hiện kháng HA của H3 (Protein Sciences Corp, Meriden, CT, USA) để thăm dò sự biểu hiện hoặc kháng thể đơn dòng đặc hiệu HA (5 microgram/ml) và kháng thể thứ cấp được sử dụng để nhuộm. Sau đó tế bào được phân tích bằng phương pháp phân loại tế bào liên kết huỳnh quang (FACS) về sự biểu hiện của polypeptit vùng gốc của HA được gắn lên màng theo sáng chế sử dụng huyết thanh đa dòng xuất hiện kháng HA của H3 hoặc HA của H1 để thăm dò sự biểu hiện. Dải kháng thể đơn dòng có tính đặc hiệu đã biết mà liên kết với protein chiều dài đủ (CR8020, CR8043 và CR9114) được sử dụng để thăm dò về sự có mặt của epitop bảo thủ và, bằng cách suy luận, sự cuộn gấp chính xác của HA chiều dài đủ và các mini-HA polypeptit theo sáng chế. Kháng thể đơn dòng CR6261 (đã biết là không liên kết với các HA của H3) và CR8057 (liên kết với vùng đầu của HA từ A/Wisconsin/67/2005) cũng được bao gồm trong thử nghiệm này. Các kết quả được biểu hiện dưới dạng tỷ lệ phần trăm tế bào dương tính và được thể hiện trong FIG. 8.

Các kết quả thể hiện rằng tất cả các cấu trúc được biểu hiện trên bề mặt tế bào phản ứng với huyết thanh đa dòng H3 dẫn đến 80-90% tất cả các tế bào được phân tích là dương tính với trình tự dựa trên H3 và nhiều hơn 50% đối với trình tự chiều dài đủ của H1 so với dưới 4% đối với tế bào không được chuyển nhiễm. Sử dụng kháng thể đa dòng kháng H1, 60-70% tất cả các tế bào là dương tính, ngoài trừ trình tự chiều dài đủ của H1 đạt tới 100%. Các thử nghiệm đối chứng không có IgG, chỉ sử dụng IgG kháng người hoặc kháng thỏ được đánh dấu tất cả đều âm tính. Cả protein HA chiều dài đủ của A/Wisconsin/67/2005 và A/Brisbane/59/2007 được nhận diện bằng kháng thể đơn dòng CR9114, đã biết là có khả năng trung hòa cả hai chủng. HA chiều dài đủ của A/Wisconsin/67/2005 còn liên kết với CR8020, CR8043, và CR8057 (chỉ liên kết với một số chủng H3, được mô tả trong WO2010/130636), nhưng không liên kết với CR6261 (Throsby et al. (2008), WO 2008/028946). Đối với HA chiều dài đủ từ

A/Brisbane/59/2007 điều ngược lại là đúng: nó liên kết với CR6261 nhưng không được nhận diện bởi CR8020, CR8043 và CR8057.

Polypeptit như được mô tả trong SEQ ID NO: 91 với trình tự SEQ ID NO: 103 không có khả năng liên kết với CR8020, CR8043 và CR9114 trong bất kỳ trường hợp nào như được chứng minh bằng việc không có các tín hiệu trên nền trong FIG. 8. Do đó kết luận rằng các trình tự này không trình diện các epitop của các kháng thể này và kết quả là các protein như có mặt trên màng tế bào không được cuộn gấp thành cấu trúc 3 chiều tự nhiên của nó.

Ví dụ 12: Thiết kế polypeptit vùng gốc chứa vùng gốc bảo thủ khác có các epitop của CR8020, CR8043 và CR9114 dựa trên HA của H3

Trong ví dụ này việc thiết kế polypeptit theo sáng chế khác dựa trên typ huyết thanh H3 được mô tả. HA của virus cúm H3 A/Wisconsin/67/2005 (SEQ ID NO: 89) và A/Hong Kong/1/1968 (SEQ ID NO: 121) được sử dụng làm trình tự bố mẹ.

Cải biến thứ nhất trong trình tự là việc loại bỏ vị trí cắt ở vị trí 345 (đánh số đề cập đến trình tự SEQ ID NO: 89 bằng cách gây đột biến R thành Q (R345Q) để ngăn chặn sự tạo thành HA1 và HA2 từ HA0. Một cách tùy ý các gốc 347 đến 351 (IFGAI, một phần của peptit dung hợp) có thể được loại bỏ thêm để làm giảm đến mức thấp nhất sự tiếp xúc của các gốc kỵ nước với dung môi chứa nước. Điện tích dương ở vị trí cắt được bảo tồn 100% trong H3 và do đó đột biến này có thể được áp dụng trong tất cả các trình tự.

Cải biến thứ hai là việc loại bỏ vùng đầu bằng cách loại bỏ một phần lớn của trình tự HA1 và nối lại các trình tự đầu tận cùng N và C thông qua một cầu nối ngắn. Đoạn bị loại bỏ có thể thay đổi về chiều dài, nhưng được ưu tiên là gốc cuối cùng của trình tự đầu tận cùng N của HA1 và gốc đầu tiên của trình tự đầu tận cùng C gần nhau về mặt không gian để tránh việc đưa chúng vào thông qua trình tự nối. Trong trình tự H3 các đoạn được loại bỏ có thể được đưa vào ở (các vị trí tương đương của) S62-P322 (mini2; SEQ ID NO: 105), S63-P305 (mini3; SEQ ID NO: 119) và T64-T317 (mini4; SEQ ID NO: 120). Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp ví dụ như Clustal hoặc Muscle. Các phần còn lại của trình tự có thể được nối trực tiếp hoặc theo cách khác có thể đưa vào cầu nối linh động. Trình tự cầu nối có thể có chiều dài từ 1 đến 50 axit amin. Được ưu tiên là các cầu nối linh động có chiều dài giới hạn (nhỏ hơn hoặc bằng 10 axit amin), ví dụ GGG, GGGG, GSA, GSAG, GSAGSA, GSAGSAG hoặc tương tự. Chiều dài của đoạn bị loại bỏ cũng có thể thay đổi, ví dụ bằng cách giảm số lượng gốc trong đoạn được loại

bỏ bằng cách bắt đầu ở (tương đương của) vị trí 63, 64, 65, 66, 67, hoặc làm tăng chiều dài của đoạn bị loại bỏ, bằng cách cắt ở vị trí 57, 58, 59, 60 hoặc 61. Tương tự, axit amin cuối cùng bị loại bỏ có thể ở (tương đương của) vị trí 317, 318, 319, 320 hoặc 321, hoặc làm tăng chiều dài của đoạn bị loại bỏ ở (tương đương của) vị trí 323, 324, 325, 326, hoặc 327. Quan trọng là nhận diện được sự thay đổi về chiều dài của đoạn được loại bỏ có thể ở trong phần được bù bằng cách làm cho phù hợp chiều dài của trình tự cầu nối, tức là đoạn được loại bỏ lớn hơn có thể được làm cho phù hợp với cầu nối dài hơn và ngược lại. Các polypeptit này cũng được bao gồm trong sáng chế.

Việc loại bỏ vùng đầu ra khỏi vòng B giữa các gốc 400 đến 420 hiện được cho tiếp xúc với dung môi chứa nước. Trong các HA của H3 vòng này có tính bảo thủ cao (xem Bảng 9). Trình tự liên ứng là: 401 I(E/G)KTNEKFHQIEKEFSEVEGR 421 (SEQ ID NO: 104; đánh số đề cập đến trình tự SEQ ID NO: 89). Để làm tăng độ hòa tan của vòng này đối với polypeptit theo sáng chế trong cấu hình trước dung hợp và làm mất ổn định cấu hình sau khi dung hợp một số gốc kỵ nước phải được biến đổi thành gốc phân cực (S,T,N,Q), axit amin tích điện (R,H,K,D,E), hoặc phải làm tăng độ linh động bằng cách gây đột biến thành G. Gây đột đặc hiệu ở các vị trí 401, 408, 411, 415, 418, (đánh số đề cập đến trình tự SEQ ID NO: 89) sẽ góp phần làm ổn định polypeptit theo sáng chế.

Đối với các vị trí F408 và F415 đột biến thành S được ưu tiên nhưng axit amin phân cực khác (T,N,Q), axit amin tích điện (R,H,K,D,E) và axit amin linh động cao (G) sẽ có tác động tương tự. Đối với vị trí 411 (I), đột biến thành T được ưu tiên; axit amin phân cực khác (S, N, Q), axit amin tích điện (R,H,K,D,E) và axit amin linh động cao (G) sẽ có tác động tương tự và do đó cũng được bao gồm trong sáng chế. Đối với vị trí 418 (V) đột biến thành G được ưu tiên. Axit amin phân cực khác (S, T, N, Q), axit amin tích điện (R, H, K, D, E) sẽ có tác động tương tự và do đó cũng được bao gồm trong sáng chế. Đối với vị trí 401 (I) đột biến thành R được ưu tiên nhưng axit amin phân cực khác (S, T, N, Q), axit amin tích điện (H, K, D, E) hoặc các gốc linh động (G) là cũng có thể được. Vì vậy polypeptit theo sáng chế chứa ít nhất một đột biến được mô tả trên đây. Các sự kết hợp của nhiều hơn một đột biến là cũng có thể được, như được thể hiện trong các trình tự SEQ ID NO 123-127 và 129-131.

Để làm ổn định cấu hình trước dung hợp của polypeptit theo sáng chế liên kết cộng hóa trị giữa hai phần xa nhau trong các trình tự bậc một nhưng gần nhau trong cấu hình được cuộn gấp trước dung hợp được đưa vào. Cuối cùng cầu disulfit được thiết kế trong polypeptit theo sáng chế, tốt hơn giữa (tương đương của) vị trí 326 và 438 trong H3 A/Wisconsin/67/2005 (SEQ ID NO: 89). Các vị trí tương đương có thể

được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp như Clustal, Muscle etc. Cầu disulfit được thiết kế được tạo ra bằng cách gây đột biến ít nhất một (nếu gốc còn lại là xystein rồi), nhưng thường là hai gốc gần nhau về không gian với xystein, mà sẽ tạo thành liên kết cộng hóa trị một cách tự phát hoặc bằng phản ứng oxi hóa hoạt hóa giữa các nguyên tử lưu huỳnh của các gốc này. Cầu xystein thay thế có thể được tạo ra giữa (tương đương của) vị trí 334 và 393 trong H3 A/Wisconsin/67/2005 (SEQ ID NO: 89) bằng cách gây đột biến các gốc này thành xystein. Trong một số trường hợp xystein ở (tương đương của) vị trí 321 được cải biến thành glyxin để tránh sự tạo thành cầu disulfit không mong muốn.

HA tự nhiên tồn tại ở dạng trime trên bề mặt tế bào. Hầu hết các tương tác giữa từng monome mà giữ dạng trime cùng nằm trong vùng đầu. Sau khi loại bỏ vùng đầu, do đó cấu trúc bậc ba được làm mất ổn định và vì vậy củng cố các tương tác giữa các monome trong phân tử bị cắt đứt sẽ làm tăng độ ổn định. Trong vùng gốc, trime hóa được trung gian bởi sự hình thành của môtip cuộn xoắn dạng trime. Bằng cách củng cố môtip này có thể tạo ra dạng trime thích hợp hơn. Trình tự liên ứng để tạo ra dạng trime xoắn cuộn, IEAIEKKIEAIEKKIEAIEKK, được đưa vào ở (tương đương của) vị trí 421 đến 441. Để tránh sự cản trở hình thành cầu disulfit giữa các vị trí 326 và 438 một trình tự thay thế ngắn hơn IEAIEKKIEAIEKKI ở (tương đương của) các vị trí 421 đến 435 cũng được sử dụng. Một cách khác là đưa vào trình tự RMKQIEDKIEEIESKQKKIEN, có nguồn gốc từ GCN4 và được biết là trime hóa, ở vị trí 421-441 hoặc trình tự ngắn hơn RMKQIEDKIEEIESK ở vị trí 421 đến 435.

Polypeptit theo sáng chế có thể chứa trình tự nội bào của HA và vùng xuyên màng vì vậy các polypeptit thu được trình diện trên bề mặt tế bào khi được biểu hiện trong tế bào. Theo các phương án khác, trình tự trong tế bào chất và trình tự xuyên màng từ (tương đương của) vị trí 522 đến đầu tận cùng C được loại bỏ do đó polypeptit được tiết ra (hòa tan được) được sản sinh sau khi biểu hiện trong tế bào. Tùy ý, một số gốc bổ sung có thể được bao gồm trong protein hòa tan được bằng cách loại bỏ trình tự từ (tương đương của) 523, 524, 525, 526, 527, 528 hoặc 529. Polypeptit hòa tan được có thể được làm ổn định hơn nữa bằng cách đưa vào trình tự đã biết để tạo cấu trúc trime, tức là AYVRKDGEWVLL (SEQ ID NO: 143) (trình tự 'cuộn gấp'), tùy ý được nối thông qua cầu nối. Cầu nối này có thể tùy ý chứa vị trí phân cắt để xử lý sau đó theo quy trình mà người có hiểu biết trung bình trong lĩnh vực đã biết. Để dễ dàng tinh chế dạng hòa tan có thể bổ sung thêm trình tự đuôi, ví dụ đuôi his (HHHHHHH) được nối thông qua cầu nối ngắn, ví dụ EGR. Theo một số phương

án cầu nối và trình tự đuôi his được thêm vào không có trình tự cuộn gấp được thể hiện.

Một gốc quan trọng của HA là vị trí 345 (Arg) bởi vì đây là vị trí trong đó xảy ra sự cắt do proteaza để làm cho sự dung hợp protein có đủ khả năng. Đột biến của Arg này thành Gln (R345Q) ngăn chặn xảy ra việc phân cắt do đó khóa protein trong trạng thái trước dung hợp.

Đột biến được mô tả trên đây được tập hợp như được mô tả dưới đây:

Cụm 9 F408S, I411T, F415S

Cụm 10V418G

Cụm 11I401R

Cụm 12K326C, S438C

Cụm 13T334C, I393C

Cụm 14 C321G

GCN4RMKQIEDKIEEIESKQKKIEN ở vị trí 421 đến 441

hoặc RMKQIEDKIEEIESK ở vị trí 421 đến 435

triIEAIEKKIEAIEKKIEAIEKK ở vị trí 421 đến 441

hoặc IEAIEKKIEAIEKKI ở các vị trí 421 đến 435

Sử dụng trình tự của HA chiều dài đủ từ H3N2 A/Wisconsin/67/2005 làm điểm khởi đầu các cụm được mô tả trên đây được kết hợp với việc loại bỏ cụm S62-P322 (mini2; SEQ ID NO: 105) để đi đến polypeptit theo sáng chế

SEQ ID NO: 105: H3-mini2

SEQ ID NO: 106: H3-mini2-cl9+10

SEQ ID NO: 107: H3-mini2-cl9+11

SEQ ID NO: 108: H3-mini2-cl9+10+11

SEQ ID NO: 109: H3-mini2-cl9+10+11-tri (trình tự tri ở vị trí 421-441)

SEQ ID NO: 110: H3-mini2-cl9+10+11-GCN4 (trình tự GCN4 ở vị trí 421-441)

SEQ ID NO: 111: H3-mini2-cl9+10+11+12

SEQ ID NO: 112: H3-mini2-cl9+10+12

SEQ ID NO: 113: H3-mini2-cl9+10+11+12-GCN4 (trình tự GCN4 ngắn ở vị trí 421-435)

SEQ ID NO: 114: H3-mini2-cl9+10+11+12-tri (trình tự tri ngắn ở vị trí 421-435)

SEQ ID NO: 115: H3-mini2-cl9+13

SEQ ID NO: 116: H3-mini2-cl9+10+11+13

SEQ ID NO: 117: H3-mini2-cl9+10+11+13-GCN4 (trình tự GCN4 ở vị trí 421- 441)

SEQ ID NO: 118: H3-mini2-cl9+10+11+13-tri (trình tự tri ở vị trí 421-441)

Ngoài ra việc loại bỏ các đoạn S63-P305 (mini3) và T64-T317 (mini4) được kết hợp với các cụm 9, 10, 11 và 14 để tạo ra polypeptit theo sáng chế

SEQ ID NO: 119: H3-mini3-cl9+10+11+12+14

SEQ ID NO: 120: H3-mini4-cl9+10+11+12+14

Sử dụng trình tự của HA chiều dài đủ từ H3N2 A/Hong Kong/1/1968 làm điểm khởi đầu các cụm được mô tả trên đây được kết hợp với việc loại bỏ cụm S62-P322 để đi đến polypeptit theo sáng chế

SEQ ID NO: 121: H3 Chiều dài đủ A/Hong Kong/1/1968

SEQ ID NO: 122: HK68 H3m2-cl9

SEQ ID NO: 123: HK68 H3m2-cl9+10

SEQ ID NO: 124: HK68 H3m2-cl9+10+11

SEQ ID NO: 125: HK68 H3m2-cl9+10+12

SEQ ID NO: 126: HK68 H3m2-cl9+10+11+12

SEQ ID NO: 127: HK68 H3m2-cl9+10+11+13

SEQ ID NO: 128: HK68 H3m2-cl9+10+11+12-tri (trình tự tri ngắn ở vị trí 421-435)

SEQ ID NO: 129: HK68 H3m2-cl9+10+11+13-tri (trình tự tri ở vị trí 421-441)

SEQ ID NO: 130: HK68 H3m2-cl9+10+11+12-GCN4 (trình tự GCN4 ngắn ở vị trí 421- 435)

SEQ ID NO: 131: HK68 H3m2-cl9+10+11+13-GCN4 (trình tự GCN4 ở vị trí 421- 441).

Các gen mã hóa trình tự protein nêu trên được tổng hợp và tách dòng vào vector biểu hiện pcADN2004 sử dụng các phương pháp thông thường mà người có hiểu biết trung bình trong lĩnh vực đã biết. Vì lý do so sánh trình tự HA chiều dài đủ của H3 A/Wisconsin/67/2005 và/hoặc H3 A/Hong Kong/1/1968 được bao gồm trong thử nghiệm.

Huyền phù tế bào HEK293F (Invitrogen) (10^6 tế bào /ml, 30 ml) được chuyển nhiễm bằng vector biểu hiện (1 μ g/ml) sử dụng 40 μ l 293transfectin làm tác nhân chuyển nhiễm và để cho tiếp tục nhân lên trong 2 ngày. Tế bào được thu gom, chia thành các phần phân ước (0,3 ml, khoảng 3×10^5 tế bào) và các phần phân ước được xử lý bằng huyết thanh đa dòng xuất hiện kháng HA của H3 (Protein Sciences Corp, Meriden, CT, USA) để thăm dò sự biểu hiện hoặc kháng thể đơn dòng đặc hiệu HA (5

microgram/ml) và kháng thể thứ cấp được sử dụng để nhuộm. Sau đó tế bào được phân tích bằng phương pháp phân loại tế bào liên kết huỳnh quang (FACS) về sự biểu hiện của polypeptit vùng gốc của HA được gắn lên màng theo sáng chế sử dụng huyết thanh đa dòng xuất hiện kháng HA của H3 hoặc HA của H1 để thăm dò sự biểu hiện. Dải kháng thể đơn dòng có tính đặc hiệu đã biết mà liên kết với protein chiều dài đủ (CR8020, CR8043 và CR9114) được sử dụng để thăm dò về sự có mặt của epitop bảo thủ và, bằng cách suy luận, sự cuộn gấp chính xác của HA chiều dài đủ và các mini-HA polypeptit theo sáng chế. Kháng thể đơn dòng CR6261 (đã biết là không liên kết với các HA của H3) và CR8057 (liên kết với vùng đầu của HA từ A/Wisconsin/67/2005) cũng được bao gồm trong các thử nghiệm này. Các kết quả được biểu hiện dưới dạng tỷ lệ phần trăm tế bào dương tính và được thể hiện trong FIG. 9 đối với HA của H3 của A/Wisconsin/67/2005 dựa trên các trình tự và FIG. 10 đối với HA của H3 của A/Hong Kong/1/1968 dựa trên các trình tự.

Các kết quả thể hiện rằng tất cả A/Wisconsin/67/2005 dựa trên các cấu trúc (FIG. 9) được biểu hiện trên bề mặt tế bào phản ứng với huyết thanh đa dòng H3 dẫn đến ca 80% hoặc lâu hơn tất cả các tế bào được phân tích là dương tính so với dưới 5% đối với tế bào không được chuyển nhiễm. Các thử nghiệm đối chứng không có IgG, chỉ sử dụng IgG kháng người hoặc kháng thỏ được đánh dấu tất cả đều âm tính. HA chiều dài đủ của A/Wisconsin/67/2005 được nhận diện bằng kháng thể đơn dòng CR8020, CR8043, CR8057 (chỉ liên kết với một số chủng H3, được mô tả trong WO2010/130636) và CR9114 được biết là có khả năng liên kết với protein này, nhưng không được nhận diện bằng mAb CR6261. Ngược lại, hầu hết polypeptit vùng gốc không được nhận diện bởi CR8020, CR8043 hoặc CR9114 với một số ngoại trừ cần lưu ý. Cá polypeptit chứa đột biến cụm 12 được nhận diện bởi CR8020 và/hoặc CR8043. H3-mini2-cl9+10+11+12 (SEQ ID NO: 111), H3-mini2-cl9+10+12 (SEQ ID NO: 112), H3-mini2-cl9+10+11+12-GCN4 (SEQ ID NO: 113) và H3-mini2-cl9+10+11+12-tri (SEQ ID NO: 114) biểu thị sự nhận diện bằng CR8020 (tỷ lệ % tế bào dương tính nằm trong khoảng từ ca 10 đến 60) và CR8043 (40 đến 70%) (như được thể hiện bằng các mũi tên). Trong số 4 cấu trúc dương tính H3-mini2-cl9+10+11+12-GCN4 (SEQ ID NO: 113) thể hiện đáp ứng lớn nhất trong thử nghiệm này. Thu được các kết quả tương tự từ cường độ phát huỳnh quang trung bình được thể hiện trong dải B (FIG. 9). H3-mini2-cl9+10+11+12 (SEQ ID NO: 111), H3-mini2-cl9+10+12 (SEQ ID NO: 112), H3-mini2-cl9+10+11+12-GCN4 (SEQ ID NO: 113) và H3-mini2-cl9+10+11+12-tri (SEQ ID NO: 114) biểu thị cường độ phát huỳnh quang trung bình tốt cao hơn nền sau khi tiếp xúc với CR8020 và CR8043 và nhuộm, với đáp

ứng cao nhất đối với H3-mini2-cl9+10+11+12-GCN4 (SEQ ID NO: 113). Không có polypeptit nào dựa trên HA từ A/Wisconsin/67/2005 có khả năng nhận diện CR9114.

Fig.10 thể hiện rằng tất cả A/Hong Kong/1/1968 dựa trên các cấu trúc (FIG. 10) được biểu hiện trên bề mặt tế bào phản ứng với huyết thanh đa dòng H3 đối với hầu hết các cấu trúc dẫn đến ca 40-60% tất cả các tế bào được phân tích là dương tính so với dưới 5% đối với tế bào không được chuyển nhiễm. Các thử nghiệm đối chứng không có IgG, chỉ sử dụng IgG kháng người hoặc kháng thỏ được đánh dấu tất cả đều âm tính. Tỷ lệ phần trăm tế bào dương tính đối với protein chiều dài đủ từ A/Hong Kong/1/1968 sau khi xử lý bằng huyết thanh đa dòng là thấp (ca 10%), nhưng thu được các tín hiệu mạnh từ liên kết của CR8020, CR8043 và CR9114 chỉ ra rằng protein có mặt trên bề mặt tế bào. CR8057 không nhận diện A/Hong Kong/1/1968 dựa trên các trình tự, chỉ protein chiều dài đủ từ A/Wisconsin/67/2005. Bốn cấu trúc (chứa đột biến cụm 12) được nhận diện bằng CR8020 và CR8043, tức là HK68 H3m2-cl9+10+12 (SEQ ID NO: 125), HK68 H3m2-cl9+10+11+12 (SEQ ID NO: 126), HK68 H3m2-cl9+10+11+12-tri (SEQ ID NO: 128 chứa trình tự tri được làm ngắn đi ở vị trí 421-435) và HK68 H3m2-cl9+10+11+12-GCN4 (SEQ ID NO: 130, chứa the trình tự GCN4 ngắn ở vị trí 421- 435), như được chỉ ra bởi % tế bào dương tính (15% hoặc cao hơn) và MFI cao hơn nên rõ rệt. Các tín hiệu mạnh nhất (MFI) thu nhận được đối với HK68 H3m2-cl9+10+11+12-GCN4 (SEQ ID NO:130); cấu trúc polypeptit vùng gốc này cũng thể hiện liên kết có thể phát hiện được với CR9114.

Kết luận là các tác giả sáng chế đã thể hiện rằng theo phương pháp được mô tả trên đây, có thể thu được polypeptit vùng gốc theo sáng chế đối với các typ huyết thanh của nhóm 2, cụ thể là virus cúm A của typ phụ H3.

Ví dụ 13: Thiết kế, biểu hiện và tinh chế một phần polypeptit vùng gốc hòa tan được chứa các epitop vùng gốc bảo thủ

Theo một số phương án, polypeptit theo sáng chế chứa trình tự nội bào và vùng xuyên màng của HA vì vậy các polypeptit thu được trình diện trên bề mặt tế bào khi được biểu hiện trong tế bào. Theo các phương án khác, trình tự trong tế bào chất và trình tự xuyên màng từ vị trí (hoặc vị trí tương đương của) 523, 524, 525, 526, 527, 528, 529 hoặc 530 đến đầu tận cùng C của HA2 (đánh số theo trình tự SEQ ID NO: 1) được loại bỏ do đó sự biểu hiện trong tế bào dẫn đến polypeptit được tiết ra (hòa tan được) có thể được sử dụng ví dụ trong vaccin. Polypeptit hòa tan được có thể được làm ổn định hơn bằng cách đưa vào trình tự đã biết để tạo cấu trúc trimer (còn được gọi là 'cuộn gấp'), tức là AYVRKDGEWVLL (SEQ ID NO: 143) tùy ý được nối thông qua cầu nối (ví dụ GSGYIPEAPRDGQAYVRKDGEWVLLSTFL). Cầu nối này có

thể tùy ý chứa vị trí phân cắt để xử lý sau khi tinh sạch theo quy trình mà người có hiểu biết trung bình trong lĩnh vực đã biết. Để dễ dàng tinh chế dạng hòa tan có thể bổ sung thêm trình tự đuôi, ví dụ đuôi histidin (sáu hoặc bảy Histidin liên tiếp) được nối thông qua cầu nối ngắn, ví dụ EGR. Theo một số phương án cầu nối và đuôi histidin được thêm vào không có trình tự cuộn gấp được thể hiện.

Theo sáng chế trình tự axit amin từ vị trí 530 của HA chiều dài đủ từ H1N1 A/Brisbane/59/2007 (đánh số theo trình tự SEQ ID NO: 1) đến gốc axit amin đầu tận cùng C của vùng HA2 được loại bỏ và được thay thế bằng các trình tự sau EGRHHHHHHH (SEQ ID NO: 81) chứa cầu nối ngắn và đuôi histidin. Sự thay đổi này được áp dụng với trình tự SEQ ID NO: 44: H1-mini2-cụm1+5+6-trim (tạo ra trình tự SEQ ID NO 144: s-H1-mini2-cụm1+5+6-trim), SEQ ID NO: 45: H1-mini2-cụm1+5+6-GCN4 (tạo ra trình tự SEQ ID NO: 145: s-H1-mini2-cụm1+5+6-GCN4), SEQ ID NO: 46: mini2-cụm1+5+6 (A/Brisbane/59/2007) (tạo ra trình tự SEQ ID NO: 146: s-H1-mini2-cụm1+5+6), SEQ ID NO: 47: mini2-cụm1+5+6 (A/Brisbane/59/2007) (tạo ra trình tự SEQ ID NO: 147: s-H1-mini2-cụm1+5+6), SEQ ID NO: 48: mini2-cụm1+5 (A/Brisbane/59/2007) (tạo ra trình tự SEQ ID NO: 148: s-H1-mini2-cụm1+5).

Tương tự, vì lý do so sánh sự thay đổi này được áp dụng với trình tự SEQ ID NO 1: H1 chiều dài đủ (A/Brisbane/59/2007) và ngoài ra vị trí cắt của HA được phục hồi bằng cách cải biến Arginin 343 thành Glutamin (đột biến R343Q) để tạo ra trình tự SEQ ID NO: 149: s-H1 chiều dài đủ R343Q). Hơn nữa hai polypeptit theo sáng chế được tạo ra có cầu nối khác giữa các phần đầu tận cùng N và đầu tận cùng C của HA1: s-H1-mini2-cụm1+5+6-nl (SEQ ID NO:150) và s-H1-mini2-cụm1+5+6-nl2 (SEQ ID NO: 151).

Các gen mã hóa trình tự protein nêu trên được tổng hợp và tách dòng vào vector biểu hiện pcADN2004neo sử dụng các phương pháp thông thường mà người có hiểu biết trung bình trong lĩnh vực đã biết. Huyền phù tế bào HEK293F (Invitrogen) được chuyển nhiễm bằng vector biểu hiện sử dụng \ 293transfectin làm tác nhân chuyển nhiễm theo quy trình đã biết rõ trong lĩnh vực và để cho tiếp tục nhân lên trong 7 ngày. Tế bào được tách từ môi trường nuôi cấy bằng cách ly tâm và loại bỏ mảnh vụn tế bào, trong khi dịch nổi chứa polypeptit hòa tan được theo sáng chế được thu gom để tiếp tục xử lý. Dịch nổi được tinh chế bằng phương pháp sắc ký ái lực kim loại cố định trên cột Ni-NTA để liên kết polypeptit theo sáng chế đã được gắn đuôi his với nhựa và thu hồi dịch chảy qua cột. Rửa cột bằng 3-10 thể tích cột 20 mM natri phosphat pH 7,4, 500 mM NaCl, 10 mM imidazol ('rửa'), 5-15 thể tích cột 20 mM natri phosphat pH

7,4, 500 mM NaCl, 100 mM imidazol ('rửa nghiêm ngặt') và rửa giải bằng 20 mM natri phosphat pH 7,4, 500 mM NaCl, 500 mM imidazol. Trong các trường hợp riêng thành phần đệm hoặc thể tích được sử dụng được điều chỉnh phù hợp để làm tăng hiệu suất hoặc độ tinh sạch, hoặc gradient tuyến tính được sử dụng thay vì gradient theo bậc. Các phân đoạn được thu gom từ đầu đến cuối và được phân tích trên SDS-PAGE và Western blot, sử dụng kháng thể đa dòng kháng HA của huyết thanh H1 để phát hiện (xem Fig.11a-h). Các kết quả cho thấy sự làm giàu rõ rệt polypeptit theo sáng chế trong dịch rửa giải so với nguyên liệu ban đầu.

Để xác nhận việc cuộn gấp chính xác và chức năng của polypeptit tinh sạch theo sáng chế, chế phẩm được thử nghiệm về liên kết của kháng thể đơn dòng CR9114. Cuối cùng kháng thể đơn dòng có khả năng gắn đuôi His (6 hoặc 7 histidin liên tiếp) ở đầu tận cùng C của protein được phủ lên đĩa 96 giếng tiêu chuẩn bằng cách cho 100 microlit dung dịch kháng thể nồng độ 1 µg/ml vào mỗi giếng và ủ qua đêm ở 4 °C. Sau khi loại bỏ dung dịch thừa và rửa, đĩa được phong bế bằng 150 microlit dung dịch sữa không kem 2% trong 1 giờ ở nhiệt độ trong phòng. Sau khi loại bỏ tác nhân phong bế và rửa, bổ sung 100 microlit dung dịch nồng độ 1 µg/ml chứa polypeptit theo sáng chế, cũng như vùng ngoài của protein chiều dài đủ tương ứng (SEQ ID NO: 149) và ủ trong 2h ở nhiệt độ trong phòng. Sau khi loại bỏ polypeptit theo sáng chế thừa, bổ sung mAb CR9114, mAb CR8020 (đối chứng âm) hoặc huyết thanh đa dòng xuất hiện kháng HA của H1 ở thỏ (đối chứng dương) ở nồng độ thay đổi giữa 2 và 20 µg/ml và ủ trong 2h ở nhiệt độ trong phòng. Liên kết được phát hiện qua kháng thể liên hợp HRP kháng người sử dụng quy trình đã biết rõ trong lĩnh vực.

Các kết quả (fig 12) thể hiện rằng kháng thể đơn dòng CR9114 có liên kết với polypeptit tinh sạch, hòa tan được theo sáng chế cũng như với vùng ngoài chiều dài đủ (Fig.12a), trong khi kháng thể đơn dòng CR8020 thì không liên kết (Fig.12b). Huyết thanh kháng thể đa dòng kháng H1 cũng liên kết với polypeptit theo sáng chế và vùng ngoài chiều dài đủ theo cách rất tương tự (Fig.12c). Do đó, kết luận rằng epitop trung hòa phổ rộng của CR9114 được bảo toàn trong polypeptit theo sáng chế, và có tính đến đặc tính không liên tục và tạo hình của epitop này, mà vùng gốc được cuộn gấp phù hợp và chấp nhận cấu hình ba chiều tương đương hoặc rất giống với cấu hình trong HA chiều dài đủ tự nhiên.

Chế phẩm polypeptit theo sáng chế không đồng nhất về kích thước như được xác định từ các kết quả SDS-PAGE và Western blot. Các tác giả sáng chế giả thiết rằng sự biến đổi này là do sự biến đổi trong mẫu glycosyl hóa của protein giữa các phân tử protein riêng biệt. Để xác nhận điều này, các phân phân ước nhỏ của chế phẩm

protein được xử lý bằng 3 đơn vị N-glycosidaza F (enzym loại bỏ các gốc carbohydrat liên kết N ra khỏi các gốc Asparagin) trong thời gian 18 giờ ở nhiệt độ 37 °C và được phân tích bằng SDS-PAGE và Western Blot. Các kết quả (Fig.13a và 13b) thể hiện rằng xử lý bằng N-glycosidaza làm tập trung các dải khuếch tán của polypeptit theo sáng chế tạo thành một dải ở khối lượng phân tử mong muốn được tính toán từ trình tự axit amin. Đây là bằng chứng rõ ràng rằng sự không đồng nhất về kích thước quan sát được thực chất là do sự biến đổi trong mẫu glycosyl hóa.

Chế phẩm polypeptit theo sáng chế còn được xác định đặc tính bằng HP-SEC. Cuối cùng khoảng 40 µg polypeptit theo sáng chế trong thể tích nằm trong khoảng 43 đến 63 µl (nồng độ Polypeptit trong khoảng 0,64 và 0,93 mg/ml) được bơm lên cột Tosoh TSK-gel G2000 SWxl được nối với thiết bị dò tán xạ ánh sáng đa góc. Các kết quả được thể hiện trên Fig.14. Đỉnh chính (thời gian lưu ca 8 phút) xuất hiện từ polypeptit theo sáng chế, và phân tách tốt từ các loài lớn hơn, chứng tỏ rằng có thể đạt được sự tinh sạch hơn nữa. Dựa trên dữ liệu của thiết bị dò tán xạ ánh sáng đa góc, các đỉnh chính tương ứng với các loại phân tử có khối lượng phân tử trong khoảng 50 đến 80 kiloDalton (xem Bảng 9) tùy thuộc vào polypeptit theo sáng chế theo nghiên cứu này. Trên cơ sở sự đa dạng và không đồng nhất về kích thước trong glycosyl hóa polypeptit được mô tả trên đây, cũng như sự phụ thuộc của các kết quả vào loại thủy động lực học của các phân tử, những con số này cần được lấy chỉ theo chỉ định.

Ví dụ 14: Biểu hiện và tinh chế một phân polypeptit vùng gốc hòa tan được chứa epitop vùng gốc bảo thủ của kháng thể đơn dòng CR9114, CR6261 và FI6v3

Để thu được chế phẩm polypeptit theo sáng chế có độ tinh khiết cao hơn, tế bào HEK293F được chuyển nhiễm bằng vector biểu hiện pcADN2004 chứa gen mã hóa s-H1-mini2-cum1+5+6-GCN4 (SEQ ID NO: 145). Người có hiểu biết trung bình trong lĩnh vực cần phải hiểu rằng trình tự dẫn đầu (hoặc trình tự tín hiệu) mà điều khiển sự vận chuyển protein trong quá trình sản sinh (tương ứng với axit amin 1-17 của trình tự SEQ ID NO: 145) sẽ không có mặt trong polypeptit cuối cùng được tiết ra. Cuối cùng, $1,0 \times 10^6$ vc/mL được cấy bằng cách ly tâm xuống (spin-down) tế bào HEK293F (Invitrogen) ở tốc độ 300g trong 5 phút và tạo huyền phù lại trong 300 mL môi trường Freestyle™ đã được làm ấm trước cho mỗi SF1000. Ủ dịch nuôi cấy trong 1 giờ ở nhiệt độ 37°C, 10% CO₂ ở 110 rpm trong tủ nuôi cấy multitron. Sau 1 giờ, AND plasmit được hút nhỏ giọt bằng pipet vào 9,9 mL môi trường Optimem để nồng độ đạt 1,0 µg/mL trong 300 mL thể tích dịch nuôi cấy. Đồng thời 440 µL 293fectin® được hút nhỏ giọt bằng pipet vào 9,9 mL môi trường Optimem và được ủ trong 5 phút ở nhiệt độ trong phòng. Sau 5 phút hỗn hợp ADN plasmit/Optimem được bổ sung vào

hỗn hợp 293fectin[®]/Optimem và ủ ở nhiệt độ trong phòng trong 20 phút. Sau khi ủ hỗn hợp ADN plasmid/293fectin[®] được bổ sung nhỏ giọt vào huyền phù tế bào. Dịch nuôi cấy được chuyển nhiễm được ủ ở nhiệt độ 37°C, 10% CO₂ và 110 rpm trong tủ nuôi cấy multitron. Vào ngày 7 tế bào được tách ra từ môi trường nuôi cấy bằng cách ly tâm (30 phút ở tốc độ 3000 g), trong khi đó dịch nổi chứa polypeptit hòa tan được theo sáng chế được lọc trên màng lọc kích thước 0,2 µm miệng chai để xử lý tiếp theo.

Để xác nhận sự có mặt của polypeptit theo sáng chế một phần phân ước nhỏ của dịch nổi như được phân tích bởi Western Blot, sử dụng kháng thể đơn dòng định hướng kháng đuôi his để thăm dò (Fig.15a). Quan sát được một số dải ở khối lượng phân tử hiện nằm trong khoảng 37 đến 50 kDa, gần hoặc cao hơn khối lượng phân tử tính toán được dựa trên axit amin của chế phẩm protein. Sự không đồng nhất về kích thước là do sự biến đổi trong mẫu glycosyl hóa, bởi vì các thử nghiệm trước đã chỉ ra rằng việc xử lý protein này bằng N-glycosidaza F để loại bỏ glycan liên kết N được gắn vào ra khỏi protein dẫn đến tập trung vào dải ở khối lượng phân tử mong muốn.

Sự có mặt của các epitop trung hòa phổ rộng trong polypeptit theo sáng chế được xác nhận bằng ELISA, sử dụng các kháng thể trung hòa phổ rộng CR6261, CR9114 và FI6v3 làm môi. Vì lý do so sánh kháng thể đơn dòng CR8020 cũng được bao gồm làm đối chứng âm trong thử nghiệm; kháng thể này có khả năng liên kết với phân tử HA từ virus nhóm 2 (ví dụ H3 và H7 HA), nhưng không liên kết với HA từ virus nhóm 1 (ví dụ H1 và H5 HA). Cuối cùng kháng thể đơn dòng có khả năng gắn đuôi His (6 hoặc 7 histidin liên tiếp) ở đầu tận cùng C của protein được phủ lên đĩa 96 giếng tiêu chuẩn bằng cách cho 100 microlit dung dịch kháng thể nồng độ 1 µg/ml vào mỗi giếng và ủ qua đêm ở 4°C. Sau khi loại bỏ dung dịch thừa và rửa, đĩa được phong bế bằng 150 microlit dung dịch sữa không kem 2% trong 1 giờ ở nhiệt độ trong phòng. Sau khi loại bỏ tác nhân phong bế và rửa, bổ sung 100 microlit dịch nổi và ủ trong 2h ở nhiệt độ trong phòng. Sau khi loại bỏ polypeptit theo sáng chế thừa, bổ sung mAb CR9114, theo dãy độ pha loãng 1:2 bắt đầu từ nồng độ 5µg/ml, và được ủ trong 2h ở nhiệt độ trong phòng. Liên kết được phát hiện qua kháng thể liên hợp HRP kháng người sử dụng quy trình đã biết rõ trong lĩnh vực. Quan sát được liên kết rõ ràng của CR9114, FI6v3 và ở mức độ kém hơn của CR6261 với polypeptit theo sáng chế, trong khi đó không quan sát được đáp ứng nào đối với CR8020 chứng tỏ rằng liên kết quan sát được là đặc hiệu với the kháng thể đơn dòng được thử nghiệm (Fig.15b).

Đối với mục đích tinh sạch 250 ml dịch nổi nuôi cấy được đưa lên cột bắt giữ His 5mL, rửa bằng 75 ml đệm rửa (20 mM TRIS, 500 mM NaCl, pH 7,8), và rửa giải bằng gradient từng bậc của imidazol (10, 50, 100, 200, 300 và 500 mM trong đệm

rửa). Sắc ký đồ (fig. 16) thể hiện nhiều đỉnh, có polypeptit theo sáng chế rửa giải ở 100 mM imidazol (đỉnh A) và 200 mM imidazol (đỉnh B). Cả hai đỉnh được thu gom, cô và đưa lên cột loại trừ theo kích thước để tinh chế thêm (Superdex 200). Thông số rửa giải được thể hiện trên Fig.17a và 17b. Các phân đoạn được thu gom và được phân tích trên SDS-PAGE (Fig.17c và d). Phân đoạn 3 có nguồn gốc từ cả hai đỉnh A và B chứa polypeptit theo sáng chế có độ tinh khiết cao. Hiệu suất cuối cùng là ca 10 µg/ml của dịch nuôi cấy. Các mẻ được tinh sạch không chứa nội độc tố (liều 5mg/kg; < 1 EU/mg); Sinh màu LAL) và tải trọng sinh học dưới 1 CFU/50µg.

Ví dụ 15. Thiết kế polypeptit vùng gốc chứa epitop vùng gốc bảo thủ của CR9114 dựa trên HA của virus cúm B

Quy trình được mô tả trên đây để thiết kế polypeptit theo sáng chế cũng được áp dụng cho virus cúm B. Trong ví dụ này, polypeptit theo sáng chế dựa trên các trình tự HA lấy từ các chủng virus thuộc cả hai giống đã biết, tức là B/Florida/4/2006 (B/Yamagata lineage) và B/Malaysia/2506/2004 (B/Victoria lineage) được mô tả. Người có hiểu biết trung bình trong lĩnh vực sẽ hiểu rằng việc sử dụng các trình tự HA khác của virus cúm B cũng có thể được bởi vì các trình tự này có tính bảo thủ tốt, cụ thể là trong vùng gốc. Do đó các polypeptit có nguồn gốc từ các trình tự HA khác của virus cúm B theo phần mô tả dưới đây cũng được bao gồm bởi sáng chế.

Cải biến thứ nhất trong trình tự HA của B/Florida/4/2006 là việc loại bỏ vị trí cắt ở vị trí 361 (đánh số đề cập đến trình tự SEQ ID NO: 132) bằng cách gây đột biến R (hoặc trong số lượng giới hạn của các trường hợp K) thành Q (R361Q) để ngăn chặn sự tạo thành HA1 và HA2 từ HA0. Một cách tùy ý các gốc 363 đến 367 (GFGAI, một phần của peptit dung hợp) có thể được loại bỏ thêm để làm giảm đến mức thấp nhất sự tiếp xúc của các gốc kỵ nước với dung môi chứa nước. Điện tích dương ở vị trí cắt được bảo toàn 100% trong HA từ virus cúm B và do đó đột biến này có thể được áp dụng trong tất cả các trình tự.

Cải biến thứ hai là việc loại bỏ vùng đầu bằng cách loại bỏ một phần lớn của trình tự HA1 và nối lại các trình tự đầu tận cùng N và C thông qua một cầu nối ngắn. Đoạn bị loại bỏ có thể thay đổi về chiều dài, nhưng được ưu tiên là gốc cuối cùng của trình tự đầu tận cùng N của HA1 và gốc đầu tiên của trình tự đầu tận cùng C gần nhau về mặt không gian để tránh việc đưa chúng vào thông qua trình tự nối. Trong các trình tự B các đoạn được loại bỏ có thể được đưa vào ở (các vị trí tương đương của) P51-I336 (m2; SEQ ID NO: 133) trong chủng B/Florida/4/2006 (SEQ ID NO:132). Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp ví dụ

như Clustal hoặc Muscle. Các phần còn lại của trình tự có thể được nối trực tiếp hoặc theo cách khác có thể đưa vào cầu nối linh động. Trình tự cầu nối có thể có chiều dài từ 1 đến 50 axit amin. Được ưu tiên là các cầu nối linh động có chiều dài giới hạn (nhỏ hơn hoặc bằng 10 axit amin), ví dụ GGG, GGGG, GSA, GSAG, GSAGSA, GSAGSAG hoặc tương tự.

SEQ ID NO: 133 mô tả polypeptit theo sáng chế chứa đoạn được loại bỏ P51-I332 (m2; SEQ ID NO: 133). Việc loại bỏ các đoạn được mô tả trên đây đảm bảo rằng các vùng không có cấu trúc được tạo thành bởi các gốc giữa P51 và N58 và giữa E306 và I337 cũng được loại bỏ; điều này có lợi đối với tính ổn định tổng thể của polypeptit theo sáng chế. Quan sát được tác động tương tự đối với polypeptit theo sáng chế có nguồn gốc từ trình tự H1 (xem trên đây).

Việc loại bỏ vùng đầu ra khỏi vòng giữa các gốc 416 đến 436 hiện được cho tiếp xúc với dung môi chứa nước. Các HA của B vòng này có tính bảo thủ cao (xem Bảng 10). Trình tự liên ứng là: LSELEVKNLQRLSGAMDELHN.

Để làm tăng độ ổn định của vòng này trong cấu hình trước dung hợp và làm mất ổn định cấu hình sau khi dung hợp một số gốc kỵ nước được cải biến thành axit amin phân cực (S,T,N,Q), axit amin tích điện (R,H,K,D,E), hoặc phải làm tăng độ linh động bằng cách gây đột biến thành G. Gây đột biến đặc hiệu ở các vị trí 421, 424, 427, 434 (đánh số đề cập đến trình tự SEQ ID NO: 132) sẽ góp phần làm ổn định polypeptit theo sáng chế.

Đối với các vị trí V421 và L427 đột biến thành T được ưu tiên nhưng axit amin phân cực khác (S,N,Q), axit amin tích điện (R,H,K,D,E) và axit amin linh động cao (G) sẽ có tác động tương tự. Đối với vị trí 424, đột biến thành S được ưu tiên. Axit amin phân cực khác (N, T, Q), axit amin tích điện (R,H,K,D,E) và axit amin linh động cao (G) sẽ có tác động tương tự. Đối với vị trí L434 đột biến thành G được ưu tiên. Axit amin phân cực khác (S, T, N, Q), axit amin tích điện (R, H, K, D, E) sẽ có tác động tương tự. Các polypeptit chứa ít nhất một đột biến được mô tả trên đây được tạo ra. Các sự kết hợp của nhiều hơn một đột biến là cũng có thể được, như được thể hiện trong các trình tự SEQ ID NO: 134-136.

Để làm ổn định cấu hình trước dung hợp của polypeptit theo sáng chế liên kết cộng hóa trị giữa hai phần xa nhau trong các trình tự bậc một nhưng gần nhau trong cấu hình được cuộn gấp trước dung hợp được đưa vào. Cuối cùng cầu disulfit được thiết kế trong polypeptit, tốt hơn là giữa (tương đương của) vị trí K340 và S454 trong HA từ B/Florida/4/2006 (SEQ ID NO: 134-136). Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp

và so sánh trình tự nhờ sử dụng thuật toán thích hợp như Clustal, Muscle etc. Cầu disulfit được thiết kế được tạo ra bằng cách gây đột biến ít nhất một (nếu gốc còn lại là xystein rồi), nhưng thường là hai gốc gần nhau về không gian với xystein, mà sẽ tạo thành liên kết cộng hóa trị một cách tự phát hoặc bằng phản ứng oxi hóa hoạt hóa giữa các nguyên tử lưu huỳnh của các gốc này.

Trong vùng gốc, trime hóa được trung gian bởi sự hình thành của môtip cuộn xoắn dạng trime. Bằng cách củng cố môtip này có thể tạo ra dạng trime thích hợp hơn. Các trình tự hỗ trợ cho việc tạo vòng cuộn xoắn trime có nguồn gốc từ GCN4 được đưa vào ở (tương đương của) vị trí 436 đến 452 RRMKQIEDKIEEILSKI (SEQ ID NO: 135), hoặc theo cách khác RRMKQIEDKIEEILSKI ở vị trí 436 đến 451 (SEQ ID NO: 136).

Quy trình tương tự cũng được áp dụng đối với HA từ B/Malaysia/2506/2004 (SEQ ID NO: 137) để tạo ra các polypeptit. So với HA từ B/Florida/4/2006, HA này có gốc asparagin bổ sung được chèn vào ở vị trí 178 như có thể được nhìn thấy dễ dàng trong khi ghép và so sánh các trình tự HA. Kết quả là vị trí cắt là ở vị trí 362, và đột biến tương ứng tạo thành ngăn chặn sự cắt ở R362Q. Việc cắt bỏ để loại vùng đầu trong trường hợp này là ví dụ ở đoạn từ P51 đến I337 (m2; SEQ ID NO: 138). Một lần nữa các phần còn lại của trình tự có thể được nối trực tiếp hoặc theo cách khác có thể đưa vào cầu nối linh động. Trình tự cầu nối có thể có chiều dài từ 1 đến 50 axit amin. Được ưu tiên là các cầu nối linh động có chiều dài giới hạn (nhỏ hơn hoặc bằng 10 axit amin), ví dụ GGG, GGGG, GSA, GSAG, GSAGSA, GSAGSAG hoặc tương tự.

SEQ ID NO:138 mô tả polypeptit chứa đoạn được loại bỏ từ P51-I332 (m2; SEQ ID NO: 138). Việc loại bỏ các đoạn được mô tả trên đây đảm bảo rằng các vùng không có cấu trúc được tạo thành bởi các gốc giữa P51 và N58 và giữa E307 và I338 cũng được loại bỏ; điều này có lợi đối với tính ổn định tổng thể của polypeptit theo sáng chế. Quan sát được tác động tương tự đối với polypeptit theo sáng chế có nguồn gốc từ trình tự H1 (xem trên đây).

Việc loại bỏ vùng đầu ra khỏi vòng giữa các gốc L420 đến H436 hiện được cho tiếp xúc với dung môi chứa nước. Trong các HA của B vòng này có tính bảo thủ cao (xem Bảng x). Trình tự liên ứng là: LSELEVKNLQRLSGAMDELHN.

Để làm tăng độ ổn định của vòng này trong cấu hình trước dung hợp và làm mất ổn định cấu hình sau khi dung hợp một số gốc kỵ nước được cải biến thành axit amin phân cực (S,T,N,Q), axit amin tích điện (R,H,K,D,E), hoặc phải làm tăng độ linh động bằng cách gây đột biến thành G. Gây đột biến đặc hiệu ở các vị trí 422, 425, 428, 435 (đánh số đề cập đến trình tự SEQ ID NO: 137) được thử nghiệm.

Đối với các vị trí V422 và L428 đột biến thành T được ưu tiên nhưng axit amin phân cực khác (S,N,Q), axit amin tích điện (R,H,K,D,E) và axit amin linh động cao (G) sẽ có tác động tương tự. Đối với vị trí 425, đột biến thành S được ưu tiên. Axit amin phân cực khác (N, T, Q), axit amin tích điện (R,H,K,D,E) và axit amin linh động cao (G) sẽ có tác động tương tự. Đối với vị trí L435 đột biến thành G được ưu tiên. Axit amin phân cực khác (S, T, N, Q), axit amin tích điện (R, H, K, D, E) sẽ có tác động tương tự. Các polypeptit chứa ít nhất một đột biến được mô tả trên đây được tạo ra. Các sự kết hợp của nhiều hơn một đột biến là cũng có thể được, như được thể hiện trong các trình tự SEQ ID NO: 139-141.

Để làm ổn định cấu hình trước dung hợp của polypeptit theo sáng chế liên kết cộng hóa trị giữa hai phần xa nhau trong các trình tự bậc một nhưng gần nhau trong cấu hình được cuộn gấp trước dung hợp được đưa vào. Cuối cùng cầu disulfit được thiết kế trong polypeptit theo sáng chế a, tốt hơn giữa (tương đương của) vị trí K341 và S455 trong HA từ B/Malaysia/2506/2004 (SEQ ID NO: 139-141). Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp như Clustal, Muscle etc. Cầu disulfit được thiết kế được tạo ra bằng cách gây đột biến ít nhất một (nếu gốc còn lại là xystein rồi), nhưng thường là hai gốc gần nhau về không gian với xystein, mà sẽ tạo thành liên kết cộng hóa trị một cách tự phát hoặc bằng phản ứng oxy hóa hoạt hóa giữa các nguyên tử lưu huỳnh của các gốc này.

Như được mô tả trên đây, HA tự nhiên tồn tại ở dạng trime trên bề mặt tế bào. Hầu hết các tương tác giữa từng monome mà giữ dạng trime cùng nằm trong vùng đầu. Sau khi loại bỏ vùng đầu, do đó cấu trúc bậc ba được làm mất ổn định và vì vậy củng cố các tương tác giữa các monome trong phân tử bị cắt đứt sẽ làm tăng độ ổn định. Trong vùng gốc, trime hóa được trung gian bởi sự hình thành của môtip cuộn xoắn dạng trime. Bằng cách củng cố môtip này có thể tạo ra dạng trime thích hợp hơn. Trình tự hỗ trợ cho việc tạo vòng cuộn xoắn trime có nguồn gốc từ GCN4 được đưa vào ở (tương đương của) vị trí 437 đến 453 RRMKQIEDKIEEILSKI (SEQ ID NO: 135), hoặc theo cách khác RRMKQIEDKIEEILSKI ở vị trí 437 đến 452 (SEQ ID NO: 136).

Polypeptit dựa trên hemagglutinin của virus cúm B, SEQ ID NO: 133-136 và 138-141 được thử nghiệm về sự có mặt của epitop của CR9114 bằng phương pháp phân loại tế bào liên kết huỳnh quang như được mô tả trên đây. Tuy nhiên, không quan sát được liên kết của mAb CR9114 đối với các cấu trúc này.

Ví dụ 16: *Khả năng sinh miễn dịch của polypeptit vùng gốc của HA thể hệ thứ*

3

Để đánh giá khả năng sinh miễn dịch của polypeptit vùng gốc các con chuột được gây miễn dịch bằng vector biểu hiện mã hóa H1 chiều dài đủ từ A/Brisbane/59/2007 (SEQ ID NO: 1), Mini3- cụm11 (SEQ ID NO: 11), Mini2- cụm11+5 (SEQ ID NO: 14), mini2- cụm1+5 (SEQ ID NO: 48), mini2- cụm1+5+6 (SEQ ID NO: 46), mini2- cụm1+5+6-GCN4 (SEQ ID NO: 45) và mini2- cụm1+5+6-nl (SEQ ID NO: 152). Vector biểu hiện mã hóa cM2 cũng được bao gồm làm đối chứng âm.

Các nhóm gồm 4 con chuột (BALB/c) được gây miễn dịch bằng 50 µg cấu trúc + 50 µg tá dược (pUMCV1-GM-CSF) tiêm bắp vào ngày 1, 21 và 42. Vào ngày 49 tiến hành lấy máu lần cuối và thu lấy huyết thanh. Huyết thanh được phân tích bằng ELISA sử dụng vùng ngoài của HA chiều dài đủ tái tổ hợp từ các chủng A/Brisbane/59/2007 và A/California/07/2009 (thu được từ hãng Protein Sciences Corporation, Meriden, CT, USA) làm kháng nguyên. Tóm lại, phủ lên đĩa 96 giếng bằng 50 ng HA qua đêm ở 4°C, sau đó ủ với đệm phong bế (100 µl PBS, pH 7,4 + 2% sữa không kem) trong 1 giờ ở nhiệt độ phòng. Rửa các đĩa bằng PBS + 0,05% Tween-20, và 100 µl của dãy độ pha loãng 2 lần trong đệm phong bế, bắt đầu từ độ pha loãng 20 lần của huyết thanh được bổ sung. Kháng thể liên kết được phát hiện bằng cách sử dụng IgG của dê kháng chuột liên hợp với HRP, sử dụng các quy trình chuẩn đã được thiết lập trong lĩnh vực này. Các độ chuẩn được so với đường cong chuẩn sử dụng mAb 3AH1 InA134 (Hytest, Turku, Finland) để nhận được số đơn vị/ml theo ELISA (EU/ml).

Tiến trình thời gian của đáp ứng IgG hướng đến vùng ngoài của protein cùng loại chiều dài đủ được gây ra bởi quy trình gây miễn dịch được mô tả trên đây được thể hiện trên Fig.18. Có thể đã quan sát được có đáp ứng cao đối với các con chuột được gây miễn dịch bằng ADN (SEQ ID NO: 1) mã hóa protein chiều dài đủ sau 4 tuần. Đáp ứng được tăng lên bằng cách tiêm tiêu tăng cường, như được thể hiện từ độ chuẩn được tăng lên ở 7 tuần. Việc gây miễn dịch bằng ADN mã hóa polypeptit theo sáng chế mini2- cụm11+5 (SEQ ID NO: 14), mini2- cụm1+5 (SEQ ID NO: 48), mini2- cụm1+5+6 (SEQ ID NO: 46), mini2- cụm1+5+6-GCN4 (SEQ ID NO: 45) và mini2- cụm1+5+6-nl (SEQ ID NO: 152) dẫn đến các độ chuẩn trung gian mà tiếp tục được tăng lên khi gây miễn dịch tăng cường như được chứng minh từ các độ chuẩn ở tuần 7. Việc gây miễn dịch bằng ADN mã hóa Mini3- cụm11 (SEQ ID NO: 11) và đối chứng âm cM2 không gây ra đáp ứng có thể phát hiện được trong thử nghiệm này.

Fig.19 thể hiện đáp ứng IgG ở tuần 7 sau lần gây miễn dịch ban đầu cho cá thể chuột kháng vùng ngoài của hemagglutinin chiều dài đủ từ chủng cùng loài H1N1 A/Brisbane/59/2007 (dải A) và chủng khác loài H1N1 A/California/07/2009. Các kháng thể được gây ra bởi ADN mã hóa polypeptit theo sáng chế ADN mã hóa polypeptit theo sáng chế mini2-cụm11+5 (SEQ ID NO: 14), mini2-cụm1+5 (SEQ ID NO: 48), mini2-cụm1+5+6 (SEQ ID NO: 46), mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) và mini2-cụm1+5+6-nl (SEQ ID NO: 152) liên kết tốt tương tự với vùng ngoài của hemagglutinin có nguồn gốc từ chủng cùng loài và khác loài. Ngược lại, việc gây miễn dịch bằng ADN mã hóa protein chiều dài đủ (SEQ ID NO: 1) dẫn đến độ chuẩn cao kháng hemagglutinin cùng loài (nhiều hơn một bậc về độ lớn so với độ chuẩn quan sát được đối với việc gây miễn dịch bằng ADN mã hóa polypeptit theo sáng chế), nhưng dẫn đến độ chuẩn thấp kháng vùng ngoài của hemagglutinin khác loài. Việc gây miễn dịch bằng ADN mã hóa Mini3-cụm11 (SEQ ID NO: 11) và đối chứng âm cM2 không gây ra đáp ứng có thể phát hiện được kháng lại các vùng ngoài của hemagglutinin trong thử nghiệm này.

Kết luận là, các kháng thể xuất hiện kháng polypeptit theo sáng chế mini2-cụm11+5 (SEQ ID NO: 14), mini2-cụm1+5 (SEQ ID NO: 48), mini2-cụm1+5+6 (SEQ ID NO: 46), mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) và mini2-cụm1+5+6-nl (SEQ ID NO: 152) có khả năng nhận diện hemagglutinin chiều dài đủ. Các epitop của chúng cần phải có mặt ở vùng gốc hemagglutinin và được bảo giữa các hemagglutinin chiều dài đủ từ H1N1 A/Brisbane/59/2007 và H1N1 A/California/07/2009.

Ví dụ 17: Khả năng sinh miễn dịch của polypeptit vùng gốc HA mini2-cụm1+5+6-GCN4 thể hệ thứ ba

Để đánh giá hơn nữa khả năng sinh miễn dịch của polypeptit vùng gốc theo sáng chế, các con chuột được gây miễn dịch bằng vector biểu hiện mã hóa mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) (khởi mào) và hai lần để tăng cường bằng protein tinh sạch s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) với khoảng cách ba tuần. Vì lý do so sánh, các nhóm tách biệt được gây miễn dịch ba lần với khoảng cách ba tuần bằng vector biểu hiện mã hóa mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) cũng như H1 chiều dài đủ từ A/Brisbane/59/2007 (SEQ ID NO: 1). Vector biểu hiện mã hóa cM2 cũng được bao gồm làm đối chứng âm.

Các nhóm gồm 4 con chuột (BALB/c) được gây miễn dịch theo đường tiêm bắp (qua đường trong cơ - i.m.) bằng 1000 µg cấu trúc mã hóa mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) + 100 µg tá dược (pUMCV1-GM-CSF) vào ngày 1 và bằng s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145; 100 µg protein tinh sạch) tá dược

bằng 10 µg Nền-M vào ngày 21 và 42. Một nhóm nhận 2nd và 3rd được gây miễn dịch qua đường dưới da, trong khi nhóm còn lại nhận 2nd và 3rd được gây miễn dịch qua đường tiêm bắp. Nhóm thứ ba lại được khơi mào bằng 100 µg cấu trúc mã hóa mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) + 100 µg tá dược (pUMCV1-GM-CSF) vào ngày 1 như trên và được gây miễn dịch tăng cường vào ngày 21 và 41 of s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145; 100 µg protein tinh sạch) có tá dược với Montanide ISA-720 (1:1 v/v). Để so sánh, các nhóm gồm 4 con chuột (BALB/c) được gây miễn dịch qua đường tiêm bắp vào ngày 1, 21 và 42 bằng 100 µg cấu trúc mã hóa mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45), H1 chiều dài đủ từ A/Brisbane/59/2007 (SEQ ID NO: 1) hoặc cM2, có tá dược với 100 µg tá dược (pUMCV1-GM-CSF).

Vào ngày 49 tiến hành lấy máu lần cuối và thu lấy huyết thanh. Huyết thanh được phân tích bằng ELISA sử dụng HA chiều dài đủ tái tổ hợp từ các chủng H1N1 A/Brisbane/59/2007, H1N1 A/California/07/2009, H5N1 A/Vietnam/1203/2004 và H3N2 A/Hong Kong//1968 (thu được từ hãng Protein Sciences Corporation, Meriden, CT, USA) làm kháng nguyên. Tóm lại, phủ lên đĩa 96 giếng bằng 50 ng HA qua đêm ở 4 °C, sau đó ủ với đệm phong bế (100 µl PBS, pH 7,4 + 2% sữa không kem) trong 1 giờ ở nhiệt độ trong phòng. Rửa các đĩa bằng PBS + 0,05% Tween-20, và 100 µl của dãy độ pha loãng 2 lần trong đệm phong bế, bắt đầu từ độ pha loãng 20 lần của huyết thanh được bổ sung. Kháng thể liên kết được phát hiện bằng cách sử dụng IgG của dê kháng chuột liên hợp với HRP, sử dụng các quy trình chuẩn đã được thiết lập trong lĩnh vực này. Các độ chuẩn được so với đường cong chuẩn gồm độ pha loãng theo dãy của kháng thể đơn dòng của chuột liên kết với kháng nguyên HA và được biểu hiện dưới dạng các đơn vị ELISA cho mỗi ml (EU/ml). Fig.20 thể hiện đáp ứng IgG ở tuần 7 sau lần gây miễn dịch ban đầu cho cá thể chuột kháng vùng ngoài của hemagglutinin chiều dài đủ từ chủng cùng loài H1N1 A/Brisbane/59/2007 (dải A), chủng khác loài H1N1 A/California/07/2009 (dải B) chủng khác typ phụ H5N1 A/Vietnam/1203/2004 (dải C) và chủng khác typ phụ H3N2 A/Hong Kong/1/1968 (dải D). Các kháng thể được sinh ra do gây miễn dịch bằng ADN mã hóa polypeptit theo sáng chế mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) có khả năng nhận diện HA của H1N1 A/Brisbane/59/2007, H1N1 A/California/07/2009 và ở mức độ kém hơn với H5N1 A/Vietnam/1203/2004. Các kháng thể được kích thích bằng việc gây miễn dịch bằng ADN mã hóa H1 chiều dài đủ từ A/Brisbane/59/2007 (SEQ ID NO: 1) nhận diện protein cùng loài rất tốt, nhưng không nhận diện HA khác loài từ H1N1 A/California/07/2009, và khác typ phụ HA từ H5N1 A/Vietnam/1203/2004 ít hơn nhiều, như được chứng minh bằng các độ chuẩn thấp trên Fig.17B và C. Nhóm các

con chuột được gây miễn dịch bằng ADN mã hóa mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45; khơi mào) sau đó, được gây miễn dịch tăng cường bằng protein s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) biểu hiện độ chuẩn cao kháng lại vùng ngoài của HA có nguồn gốc từ H1N1 cùng loài A/Brisbane/59/2007, H1N1 khác loài A/California/07/2009 và H5N1 khác typ phụ A/Vietnam/1203/2004.

Fig.20D thể hiện đáp ứng IgG vào tuần 7 kháng vùng ngoài của HA từ H3N2 A/Hong Kong/1/1968. không giống với mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) và s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) có nguồn gốc từ HA of H1N1 A/Brisbane/59/2007, chủng thuộc virus cúm nhóm 1, H3N2 A/Hong Kong/1/1968 thuộc virus cúm nhóm 2 và do đó trình tự bố mẹ khác nhau về mặt phát sinh chủng loài được sử dụng để thiết kế polypeptit theo sáng chế được sử dụng trong thử nghiệm này. Gây miễn dịch ba lần bằng ADN mã hóa mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) hoặc HA chiều dài đủ từ H1N1 A/Brisbane/59/2007 (SEQ ID NO: 1) không dẫn đến các mức IgG có thể phát hiện được bằng ELISA kháng kháng nguyên này. Ngược lại, việc gây miễn dịch bằng ADN mã hóa mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) sau đó gây miễn dịch hai lần để tăng cường bằng protein tinh sạch s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) dẫn đến độ chuẩn cao kháng HA từ H3N2 A/Hong Kong/1/1968. Kết quả này thu được không phụ thuộc vào đường gây miễn dịch được sử dụng (tức là tiêm bắp và dưới da) hoặc the tá được được bổ sung vào việc gây miễn dịch tăng cường bằng protein (Matrix-M hoặc Montanide ISA-720).

Kết luận là, việc gây miễn dịch bằng polypeptit theo sáng chế có thể kích thích các IgG mà có khả năng nhận diện HA từ phổ rộng các chủng virus cúm, bao gồm các chủng cùng loài, khác loài và khác typ phụ từ virus cúm 1 cũng như chủng từ virus cúm 2. Ngược lại việc gây miễn dịch bằng HA chiều dài đủ dẫn đến độ chuẩn cao kháng HA của chủng cùng loài, làm giảm độ chuẩn của chủng kháng khác ngoài và khác typ phụ và các mức IgG thấp hơn giới hạn phát hiện đối với chủng từ virus cúm 2.

Ví dụ 18: *Thiết kế polypeptit vùng gốc chứa vùng gốc bảo thủ khác có các epitop của CR6261 và CR9114*

Polypeptit theo sáng chế được thiết kế theo quy trình được mô tả trên đây có thể được cải biến hơn nữa để làm tăng độ ổn định. Các cải biến này có thể được đưa vào để tăng cường sự hình thành các dạng trime của polypeptit theo sáng chế so với các loại monome và/hoặc đime. Như được mô tả, HA tự nhiên tồn tại ở dạng trime trên bề mặt tế bào. Nhiều tương tác giữa từng monome mà giữ dạng trime cùng nằm trong vùng đầu. Sau khi loại bỏ vùng đầu, do đó cấu trúc bậc ba được làm mất ổn định

và vì vậy củng cố các tương tác giữa các monome trong phân tử bị cắt cụt sẽ làm tăng độ ổn định của dạng trime. quá trình trime hóa được trung gian bởi sự hình thành dạng trime. Bằng cách củng cố môtip vòng cuộn xoắn trong vùng gốc dạng trime ổn định hơn có thể đạt được.

Theo sáng chế, trình tự liên ứng để tạo ra dạng trime xoắn cuộn, IEAIEKKIEAIEKKIE (SEQ ID NO: 83), được đưa vào trong polypeptit theo sáng chế ở (tương đương của) vị trí 418 đến 433 (SEQ ID NO: 44) trong chủng H1 A/Brisbane/59/2007 (đánh số theo trình tự SEQ ID NO: 1). Theo cách khác IEAIEKKIEAIEKKI (SEQ ID NO: 85) có thể được đưa vào ở 419-433 (SEQ ID NO: 49) hoặc IEAIEKKIEAIEKK (SEQ ID NO: 86) ở vị trí 420-433 (SEQ ID NO: 50). Một cách khác là đưa vào trình tự MKQIEDKIEEIESKQ (SEQ ID NO: 84), có nguồn gốc từ GCN4 và được biết là trime hóa, ở vị trí 419-433 (SEQ ID NO: 45). Theo cách khác MKQIEDKIEEIESK (SEQ ID NO: 87) có thể được đưa vào ở vị trí 420-433 (SEQ ID NO: 51) hoặc RMKQIEDKIEEIESKQK (SEQ ID NO: 88) ở vị trí 417-433 (SEQ ID NO: 52). Tương tự, bề mặt trung gian của trime có thể được gia cố bằng cách cải biến M420, L423, V427, G430 thành isoleuxin (SEQ ID NO: 53).

Theo một số phương án, polypeptit theo sáng chế chứa trình tự nội bào của HA của H1 và vùng xuyên màng. Theo các phương án khác, trình tự trong tế bào chất và trình tự xuyên màng từ vị trí (hoặc dạng tương đương của nó) 523, 524, 525, 526, 526, 527, 528, 529, hoặc 530 của HA2 đến đầu tận cùng C của HA2 (đánh số theo trình tự SEQ ID NO: 1) được loại bỏ do đó polypeptit được tiết ra (hòa tan được) được sản sinh. Polypeptit hòa tan được có thể được làm ổn định hơn như được mô tả trên đây.

Mô tả các biến thể cần nổi

Các gen mã hóa trình tự protein nêu trên (SEQ ID NO: 44 đến 46; SEQ ID NO: 49 đến 53 và SEQ ID NO: 152-157 được tổng hợp và tách dòng vào vector biểu hiện pcADN2004 sử dụng các phương pháp thông thường mà người có hiểu biết trung bình trong lĩnh vực đã biết. Vì lý do so sánh vector biểu hiện mã hóa trình tự chiều dài đủ (SEQ ID NO: 1) cũng như cM2 được bao gồm trong thử nghiệm.

Huyền phù tế bào HEK293F (Invitrogen) (10^6 tế bào/ml, 30 ml) được chuyển nhiễm bằng vector biểu hiện (1 μ g/ml) sử dụng 40 μ l 293-transfectin làm tác nhân chuyển nhiễm và để cho tiếp tục nhân lên trong 2 ngày. Tế bào được thu gom; chia thành các phần phân ước (0,3 ml, khoảng 3×10^5 tế bào) và các phần phân ước được xử lý bằng huyết thanh đa dòng xuất hiện kháng HA của H1 để thăm dò sự biểu hiện hoặc kháng thể đơn dòng đặc hiệu HA (5 microgram/ml) và kháng thể thứ cấp được sử dụng để nhuộm. Sau đó tế bào được phân tích bằng phương pháp phân loại tế bào liên

kết huỳnh quang (FACS) về sự biểu hiện của polypeptit vùng gốc của HA được gắn lên màng theo sáng chế sử dụng huyết thanh đa dòng xuất hiện kháng HA của H1 để thăm dò sự biểu hiện. Dải kháng thể đơn dòng có tính đặc hiệu đã biết mà liên kết với protein chiều dài đủ (CR6261, CR9114, CR9020 và CR8020) được sử dụng để thăm dò về sự có mặt của epitop bảo thủ và, bằng cách suy luận, sự cuộn gấp chính xác của HA chiều dài đủ và các mini-HA polypeptit theo sáng chế. Các kết quả được biểu hiện dưới dạng tỷ lệ phần trăm tế bào dương tính và cường độ phát huỳnh quang trung bình và được thể hiện trong FIG 21.

Các kết quả thể hiện rằng tất cả các biến thể được thử nghiệm được biểu hiện trên bề mặt tế bào như được chứng minh bằng đáp ứng dương tính từ huyết thanh kháng thể đa dòng kháng H1. Kháng thể CR8020 đặc hiệu HA của H3 không nhận diện bất kỳ cấu trúc nào được bao gồm trong thử nghiệm, trong khi CR9020, liên kết với vùng đầu của các HA của H1 chỉ nhận diện rõ ràng protein chiều dài đủ. Tất cả polypeptit theo sáng chế, cũng như protein chiều dài đủ được nhận diện bằng CR6261 và CR9114, chứng tỏ rằng các epitop tương ứng có mặt trong polypeptit theo sáng chế trong cấu hình tương tự như trong protein kiểu đại. Trong số polypeptit theo sáng chế có môtip trime hóa bổ sung được bao gồm trong chuỗi cuộn xoắn CD (xem FIG. 1), SEQ ID NO: 45, 51 và 52 chứa các trình tự có nguồn gốc từ GCN4 SEQ ID NO: 84, 87 và 88, tương ứng dẫn đến đáp ứng tương tự hoặc cao hơn (MFI) so với SEQ ID NO: 44, 49 và 50, chứa các trình tự trime hóa liên ứng của trình tự SEQ ID NO: 83, 85 và 86.

Sự biến đổi trong thành phần của cầu nối nối axit amin 52 và 321 (đánh số đề cập đến trình tự SEQ ID NO: 1) trong polypeptit theo sáng chế không dẫn đến sự thay đổi chính trong việc nhận diện của kháng thể đơn dòng CR6261 và CR9114. Quan sát được sự thay đổi lớn nhất khi GGGG trong SEQ ID NO: 46 được thay thế bằng HNGK, tạo ra trình tự SEQ ID NO: 152, dẫn đến đáp ứng ở mức độ thấp hơn với CR6261, nhưng không tác động đến đáp ứng với CR9114. Loại bỏ cầu nối và đưa các axit amin 53-56 có trình tự SEQ ID NO: 1 (SHNG), tức là tạo polypeptit theo sáng chế không có cầu nối trong SEQ ID NO: 46 (tạo ra trình tự SEQ ID NO: 153), SEQ ID NO: 45 (tạo ra trình tự SEQ ID NO: 154) hoặc SEQ ID NO: 50 (tạo ra trình tự SEQ ID NO: 155) không tác động đến đáp ứng trong thử nghiệm FACS, chứng tỏ rằng trình tự cầu nối không phải yếu tố quyết định.

SEQ ID NO: 156 có nguồn gốc từ SEQ ID NO: 46 bằng cách đưa vào đột biến I337N, I340N và F352Y, trong khi SEQ ID NO: 157 chứa đột biến bổ sung ở vị trí

353, tức là I353N. Các đột biến này không dẫn đến cải thiện đáp ứng với CR6261 và CR9114 trong thử nghiệm FACS được thể hiện trong FIG. 21.

Theo một số phương án, polypeptit theo sáng chế chứa trình tự nội bào của HA và vùng xuyên màng. Theo các phương án khác, trình tự trong tế bào chất và trình tự xuyên màng từ vị trí (hoặc dạng tương đương của nó) 523, 524, 525, 526, 526, 527, 528, 529, hoặc 530 của HA2 đến đầu tận cùng C của HA2 (đánh số theo trình tự SEQ ID NO: 1) được loại bỏ, và tùy ý được thay thế bằng cách đưa trình tự đã biết để tạo cấu trúc trime, tức là AYVRKDGGEWVLL (SEQ ID NO: 143), tùy ý được nối thông qua cầu nối. Cầu nối này có thể tùy ý chứa vị trí phân cắt để xử lý sau đó theo quy trình mà người có hiểu biết trung bình trong lĩnh vực đã biết. Để dễ dàng tinh chế dạng hòa tan có thể bổ sung thêm trình tự đuôi, ví dụ đuôi his HHHHHHHH được nối thông qua cầu nối ngắn, ví dụ EGR. Theo sáng chế, trình tự axit amin từ vị trí 530 (đánh số theo trình tự SEQ ID NO: 1) đến gốc axit amin đầu tận cùng C của vùng HA2 được loại bỏ và được thay thế bằng SEQ ID NO: 81 hoặc SEQ ID NO: 82.

Ví dụ 19: Khả năng sinh miễn dịch của polypeptit vùng gốc của HA thể hệ thứ

3

Để đánh giá khả năng sinh miễn dịch của polypeptit vùng gốc các con chuột được gây miễn dịch bằng vector biểu hiện mã hóa H1 chiều dài đủ từ A/Brisbane/59/2007 (SEQ ID NO: 1), Mini3- cụm11 (SEQ ID NO: 11), Mini2- cụm11+5 (SEQ ID NO: 14), mini2- cụm1+5+6 (SEQ ID NO: 46), mini2- cụm1+5+6- GCN4 (SEQ ID NO: 45), mini2- cụm1+5+6-nl (SEQ ID NO: 152), mini2- cụm1+5+6-nl2 (SEQ ID NO: 153), mini2- cụm1+5+6-nl2s-GCN4 (SEQ ID NO: 154), mini2- cụm1+5+6-GCN4t2 (SEQ ID NO: 51), mini2- cụm1+5+6-GCN4t3 (SEQ ID NO: 52), mini2- cụm1+5+6+12 (SEQ ID NO: 156) và mini2- cụm1+5+6+12+13 (SEQ ID NO: 157). Vector biểu hiện mã hóa cM2 cũng được bao gồm làm đối chứng âm.

Các nhóm gồm 4 con chuột (BALB/c) được gây miễn dịch bằng 100 µg cấu trúc + 100 µg tá dược (pUMCV1-GM-CSF) tiêm bắp vào ngày 1, 21 và 42. Vào ngày 49 tiến hành lấy máu lần cuối và thu lấy huyết thanh. Huyết thanh được phân tích bằng ELISA sử dụng HA chiều dài đủ tái tổ hợp từ các chủng H1N1 A/Brisbane/59/2007, H1N1 A/California/07/2009 và H5N1 A/Vietnam/1203/2004 (thu được từ Protein Sciences Corporation, Meriden, CT, USA) làm kháng nguyên. Tóm lại, phủ lên đĩa 96 giếng bằng 50 ng HA qua đêm ở 4 °C, sau đó ủ với đệm phong bế (100 µl PBS, pH 7,4 + 2% sữa không kem) trong 1 giờ ở nhiệt độ phòng. Rửa các đĩa bằng PBS + 0,05% Tween-20, và 100 µl của dãy độ pha loãng 2 lần trong đệm phong bế, bắt đầu từ độ pha loãng 50 lần của huyết thanh được bổ sung. Kháng thể liên kết được phát

hiện bằng cách sử dụng IgG của dê kháng chuột liên hợp với HRP, sử dụng các quy trình chuẩn đã được thiết lập trong lĩnh vực này. Các độ chuẩn được so với đường cong chuẩn gồm độ pha loãng theo dãy của kháng thể đơn dòng của chuột liên kết với kháng nguyên HA và được biểu hiện dưới dạng các đơn vị ELISA cho mỗi ml (EU/ml).

Fig.22 thể hiện đáp ứng IgG ở tuần 7 sau lần gây miễn dịch ban đầu cho cá thể chuột kháng hemagglutinin chiều dài đủ từ chủng cùng loài H1N1 A/Brisbane/59/2007 (dải A), chủng khác loài H1N1 A/California/07/2009 (dải B) và chủng khác typ phụ H5N1 A/Vietnam/1203/2004 (dải C). Các kháng thể được sinh ra do gây miễn dịch bằng ADN mã hóa polypeptit theo sáng chế Mini2-cụm11+5 (SEQ ID NO: 14), mini2-cụm1+5+6 (SEQ ID NO: 46), mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45), mini2-cụm1+5+6-nl (SEQ ID NO: 152), mini2-cụm1+5+6-nl2 (SEQ ID NO: 153), mini2-cụm1+5+6-nl2s-GCN4 (SEQ ID NO: 154), mini2-cụm1+5+6-GCN4t2 (SEQ ID NO: 51), mini2-cụm1+5+6-GCN4t3 (SEQ ID NO: 52), mini2-cụm1+5+6+12 (SEQ ID NO: 156) và mini2-cụm1+5+6+12+13 (SEQ ID NO: 157) liên kết tốt tương tự với vùng ngoài của hemagglutinin có nguồn gốc từ chủng cùng loài H1N1 A/Brisbane/59/2007 và chủng khác loài H 1N1 A/California/07/2009 (Fig.22A và B). Quan sát được độ chuẩn cao nhất đối với mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45), mini2-cụm1+5+6-GCN4t2 (SEQ ID NO: 51) và mini2-cụm1+5+6-nl2s-GCN4 (SEQ ID NO: 154). Ngược lại, việc gây miễn dịch bằng ADN mã hóa protein chiều dài đủ (SEQ ID NO: 1) dẫn đến độ chuẩn cao kháng hemagglutinin cùng loài (nhiều hơn một bậc về độ lớn so với độ chuẩn quan sát được đối với việc gây miễn dịch bằng ADN mã hóa polypeptit theo sáng chế), nhưng ở độ chuẩn thấp kháng vùng ngoài của hemagglutinin khác loài (nhiều hơn một bậc về độ lớn). Việc gây miễn dịch bằng ADN mã hóa Mini3-cụm11 (SEQ ID NO: 11) và đối chứng âm cM2 không gây ra đáp ứng có thể phát hiện được kháng các vùng ngoài của hemagglutinin trong thử nghiệm này.

Các độ chuẩn kháng vùng ngoài của hemagglutinin khác typ phụ từ H5N1 A/Vietnam/1203/2004 (Fig.22C) cho thấy đáp ứng rõ rệt đối với mini2-cụm1+5+6-GCN4t2 (SEQ ID NO: 51). Cũng thu được độ chuẩn có thể quan sát được đối với 2 trong số 4 con chuột sau khi gây miễn dịch bằng ADN mã hóa mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45) và đối với 1 trong số 4 con chuột đối với mini2-cụm1+5+6-nl (SEQ ID NO: 152), mini2-cụm1+5+6-nl2 (SEQ ID NO: 153), mini2-cụm1+5+6-nl2s-GCN4 (SEQ ID NO: 154). Ngạc nhiên là, các tác giả sáng chế cũng phát hiện được các độ chuẩn có thể phát hiện được sau khi gây miễn dịch bằng ADN mã hóa Mini3-cụm11 (SEQ ID NO 11) và Mini2-cụm11+5 (SEQ ID NO: 14). Cấu trúc cũ không gây

ra độ chuẩn kháng thể có thể phát hiện được kháng HA cùng loài và khác loài của H1, trong khi cấu trúc sau chỉ điều biến đáp ứng (Fig.22A và B). So sánh các trình tự của tất cả cấu trúc trong thử nghiệm này và HA của H5 chỉ hướng đến epitop thẳng giả định nằm ở đầu xa mang của chuỗi cuộn xoắn CD dài (xem Fig.1). Đột biến metionin thành isoleuxin ở vị trí 175 trong trình tự SEQ ID NO: 11 và vị trí 156 trong trình tự SEQ ID NO: 14 dẫn đến trình tự thẳng ERRIENLNKK (vị trí 172 đến 181 trong trình tự SEQ ID NO: 11; vị trí 153 to 162 trong trình tự SEQ ID NO: 14). Trình tự này cũng có mặt trong HA từ H5N1 A/Vietnam/1203/2004, nhưng không có trong HA từ H1N1 A/Brisbane/59/2007 và H1N1 A/California/07/2009, trong đó các trình tự tương ứng lần lượt là ERRMENLNKK và EKRIENLNKK.

Kết luận là, các kháng thể xuất hiện kháng polypeptit theo sáng chế mini2-cụm11+5 (SEQ ID NO: 14), mini2-cụm1+5+6 (SEQ ID NO: 46), mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45), mini2-cụm1+5+6-nl (SEQ ID NO: 152), mini2-cụm1+5+6-nl2 (SEQ ID NO: 153), mini2-cụm1+5+6-nl2s-GCN4 (SEQ ID NO: 154), mini2-cụm1+5+6-GCN4t3 (SEQ ID NO: 52), mini2-cụm1+5+6-GCN4t2 (SEQ ID NO: 51), mini2-cụm1+5+6+12 (SEQ ID NO: 156) và mini2-cụm1+5+6+12+13 (SEQ ID NO: 157) có khả năng nhận diện hemagglutinin chiều dài đủ. Các epitop của các kháng thể này phải có mặt ở vùng gốc hemagglutinin và được bảo giữ các hemagglutinin chiều dài đủ từ H1N1 A/Brisbane/59/2007 và H1N1 A/California/07/2009. Các kháng thể được sản sinh ra thông qua việc gây miễn dịch bằng ADN mã hóa mini2-cụm1+5+6-GCN4t2 (SEQ ID NO: 51), và ở mức độ kém hơn với mini2-cụm11+5 (SEQ ID NO: 14), mini2-cụm1+5+6-GCN4 (SEQ ID NO: 45), mini2-cụm1+5+6-nl (SEQ ID NO: 152), mini2-cụm1+5+6-nl2 (SEQ ID NO: 153), mini2-cụm1+5+6-nl2s-GCN4 (SEQ ID NO: 154) và mini2-cụm1+5+6+12 (SEQ ID NO: 156) cũng có thể nhận diện vùng ngoài của HA từ H5N1 A/Vietnam/1203/2004. Polypeptit theo sáng chế Mini3-cụm11 (SEQ ID NO: 11) có khả năng gây cảm ứng các kháng thể mà nhận diện từ H5N1 A/Vietnam/1203/2004.

Ví dụ 20. Phương pháp phổ biến để thiết kế polypeptit vùng gốc chứa vùng gốc bảo thủ có các epitop của CR6261 và CR9114

Dựa trên các kết quả được mô tả trên đây, phương pháp thông thường được xác định để tạo ra polypeptit theo sáng chế từ trình tự HA0 của virus cúm, cụ thể là từ trình tự HA0 của virus cúm thuộc typ huyết thanh H1. Phương pháp này bao gồm các bước:

1. Loại bỏ vị trí cắt giữa HA1 và HA2. Bước này có thể đạt được bằng cách gây đột biến R (một số ít trường hợp là K) thành Q ở vị trí P1 (xem ví dụ Sun et al, 2010 để biết giải thích về danh pháp của vị trí cắt (vị trí 343 trong trình tự SEQ ID

NO: 1). Gây đột biến thành Q được ưu tiên nhưng S, T, N, D hoặc E là các dạng thay thế.

2. Loại bỏ vùng đầu bằng cách loại bỏ axit amin 53 đến 320 từ trình tự SEQ ID NO: 1, hoặc ở các vị trí tương đương trong HA từ virus cúm khác. Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự sử dụng thuật toán thích hợp ví dụ như Clustal hoặc Muscle. Các phần còn lại của trình tự có thể được nối trực tiếp hoặc theo cách khác bằng cách đưa vào một cầu nối linh động. Trình tự cầu nối có thể có chiều dài từ 1 đến 50 axit amin. Được ưu tiên là các cầu nối linh động có chiều dài giới hạn (nhỏ hơn hoặc bằng 10 axit amin), ví dụ GGG, GGGG, GSA, GSAG, GSAGSA, GSAGSAG hoặc tương tự. Chiều dài của đoạn bị loại bỏ cũng có thể thay đổi, ví dụ bằng cách bắt đầu loại bỏ ở (tương đương của) vị trí 54, 55, 56, 57 hoặc 58, hoặc làm tăng chiều dài của đoạn bị loại bỏ, bằng cách cắt ở vị trí 47, 48, 49, 50, 51, hoặc 52. Tương tự, axit amin cuối cùng bị loại bỏ có thể ở (tương đương của) vị trí 315, 316, 317, 318 hoặc 319, hoặc làm tăng chiều dài của đoạn bị loại bỏ ở (tương đương của) vị trí 321, 322, 323, 324, hoặc 325. Quan trọng là nhận diện được sự thay đổi về chiều dài của đoạn được loại bỏ có thể ở trong phần được bù bằng cách làm cho phù hợp chiều dài của trình tự cầu nối, tức là đoạn được loại bỏ lớn hơn có thể được làm cho phù hợp với cầu nối dài hơn và ngược lại. Các polypeptit này cũng có thể được bao gồm trong sáng chế.

3. Làm tăng độ hòa tan của vòng (giữa chuỗi cuộn xoắn A và chuỗi cuộn xoắn CD) được tạo thành bởi (tương đương của) các gốc 402 đến 418 trong chủng H1 A/Brisbane/59/2007 (SEQ ID NO: 1) để làm tăng độ ổn định của cấu hình trước khi dung hợp và làm mất ổn định cấu hình sau khi dung hợp of the modified HA. Vòng này có tính bảo thủ cao trong trình tự H1, như được thể hiện trong bảng 6 dưới đây. Điều này có thể đạt được chẳng hạn bằng cách thay thế I, L, F hoặc V các gốc trong vòng này bằng các bản sao ưa nước. Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp ví dụ như Clustal hoặc Muscle. Đột biến thành glyxin làm mất ổn định cấu hình sau khi dung hợp do độ linh động cao của axit amin này dẫn đến sự giảm độ ổn định của chuỗi cuộn xoắn sau dung hợp được tạo thành bởi phần này của trình tự HA. Trình tự liên ứng mô tả vòng giữa gốc 402-418 của HA của virus cúm thuộc typ huyết thanh H1 là (SEQ ID NO: 17) MNTQFTAVGKEFN(H/K)LE(K/R). Trong polypeptit theo sáng chế axit amin ở các vị trí 406, 409, 413 và/hoặc 416 (hoặc vị trí tương đương của chúng, như được xác

định từ sắp xếp và so sánh trình tự) là axit amin phân cực (S,T,N,Q), tích điện (R,H,K,D,E) hoặc linh động (G). Cần lưu ý rằng đột biến của L416 thành S hoặc T cũng đưa vào vị trí glycosyl hóa N liên ứng (trình tự liên ứng là NX(S/T)). Glycosyl hóa Asparagin ở vị trí này sẽ làm tăng hơn nữa độ ổn định của vùng này. Cũng có thể là tổ hợp của các đột biến này ở các vị trí này, ví dụ F406S, V409T, L416S như trong trình tự SEQ ID NO: 10 và SEQ ID NO: 14. Trong một số trường hợp đột biến để phục hồi axit amin liên ứng được ưu tiên, ví dụ trong đó V hoặc M ở vị trí 404 (thành T), V ở 408 (thành A) hoặc 410 (thành G) hoặc I ở 414 (thành N); tỷ lệ đột biến trong trình tự ở các axit amin này là rất thấp. Tổng quát về đột biến được mô tả trên đây mà xác định đặc tính polypeptit theo sáng chế được đưa ra trong bảng 6.

4. Đưa cầu disulfit vào trong polypeptit theo sáng chế, tốt hơn là giữa axit amin ở (tương đương của) vị trí 324 và 436 trong chủng H1 A/Brisbane/59/2007; SEQ ID NO: 13-16. Các vị trí tương đương có thể được xác định dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực bằng cách sắp xếp và so sánh trình tự nhờ sử dụng thuật toán thích hợp như Clustal, Muscle etc. Cầu disulfit được thiết kế được tạo ra bằng cách gây đột biến ít nhất một (nếu gốc còn lại là xystein rồi), nhưng thường là hai gốc gần nhau về không gian với a xystein, mà sẽ tạo thành liên kết cộng hóa trị một cách tự phát hoặc bằng phản ứng oxi hóa hoạt hóa giữa các nguyên tử lưu huỳnh của các gốc này.

Sử dụng phương pháp thông thường theo sáng chế, được mô tả trên đây, polypeptit theo sáng chế được tạo ra dựa trên các trình tự HA0 của H1N1 A/California/04/2009 (SEQ ID NO: 159), H1N1 A/California/07/2009 (SEQ ID NO: 56), H1N1 A/Puerto Rico/8/1934 (SEQ ID NO: 78), và H1N1 A/Texas/36/1991 (SEQ ID NO: 64). Ngoài ra phương pháp được áp dụng cho HA từ typ phụ khác mà là một phần của Nhóm 1, tức là H5, sử dụng HA từ H5N1 A/Vietnam/1203/2004 (SEQ ID NO: 158).

H1 mini-HA A/California/07/2009 (SEQ ID NO: 160) được tạo ra từ H1 FL HA A/California/07/2009 (SEQ ID NO: 56) bằng cách

1. Loại bỏ vị trí cắt: đột biến R344Q (đánh số đề cập đến trình tự SEQ ID NO: 56)
2. Loại bỏ các gốc K53 đến P321 và đưa cầu nối GGGG giữa D52 và K322 (đánh số đề cập đến trình tự SEQ ID NO: 56)
3. Đưa gốc Serin vào vị trí 407, 417 (F407S, L417S; đánh số đề cập đến trình tự SEQ ID NO: 2), Treonin vào vị trí 410 (V410T; đánh số đề cập đến trình tự SEQ ID NO: 56) và gốc Glyxin ở vị trí 414 (F414G; đánh số đề cập đến trình tự SEQ

ID NO: 2) trong vòng giữa chuỗi cuộn xoắn A và chuỗi cuộn xoắn CD (các gốc 403-419 trong trình tự SEQ ID NO: 56)

4. Đưa cầu disulfit vào bằng cách gây đột biến các gốc Lysin 325 và Treonin 437 thành xystein (K325C, T437C; đánh số đề cập đến trình tự SEQ ID NO: 56)

5. Yếu tố làm ổn định bổ sung được đưa vào bằng cách thay thế 419KRIENLNKKVDDGFLD434 (đánh số đề cập đến trình tự SEQ ID NO: 56) bằng trình tự RMKQIEDKIEEIESKQ.

Trình tự mini-HA dựa trên HA chiều dài đủ từ A/California/04/2009 (SEQ ID NO: 159) có thể được tạo ra theo cách tương tự và giống với trình tự của H1 mini-HA A/California/07/2009 (SEQ ID NO: 160).

Tương tự, H1 mini-HA A/Puerto Rico/8/1934 (SEQ ID NO: 161) được tạo ra từ H1 FL HA A/Puerto Rico/8/1934 (SEQ ID NO: 78) bằng cách

1. Loại bỏ vị trí cắt: đột biến R343Q (đánh số đề cập đến trình tự SEQ ID NO: 78)

2. Loại bỏ các gốc S53 đến P320 và đưa cầu nối GGGG vào giữa D52 và K321 (đánh số đề cập đến trình tự SEQ ID NO: 78)

3. Đưa gốc Serin vào vị trí 406, 416 (F406S, L416S; đánh số đề cập đến trình tự SEQ ID NO: 78) Treonin vào vị trí 409 (V409T; đánh số đề cập đến trình tự SEQ ID NO: 78) và gốc Glyxin vào vị trí 413 (F413G; đánh số đề cập đến trình tự SEQ ID NO: 78) trong vòng giữa chuỗi cuộn xoắn A và chuỗi cuộn xoắn CD (các gốc 402-418 trong trình tự SEQ ID NO: 78)

4. Đưa cầu disulfit vào bằng cách gây đột biến các gốc Arginin 324 và Treonin 436 thành xystein (R324C, T436C; đánh số đề cập đến trình tự SEQ ID NO: 78)

5. Yếu tố làm ổn định bổ sung được đưa vào bằng cách thay thế 418KR MENLNKKVDDGFLD433 (đánh số đề cập đến trình tự SEQ ID NO: 78) bằng trình tự RMKQIEDKIEEIESKQ.

Sự khác nhau nữa giữa H1 mini-HA A/Puerto Rico/8/1934 (SEQ ID NO: 161) và H1 FL HA A/Puerto Rico/8/1934 (SEQ ID NO: 78) là ở vị trí 397, vị trí này là Serin trong protein chiều dài đủ (SEQ ID NO: 78) nhưng là Treonin trong polypeptit theo sáng chế của trình tự SEQ ID NO: 161 (S397T đột biến). Đây là sự biến đổi có trong tự nhiên trong trình tự A/Puerto Rico/8/1934, và các trình tự chứa đột biến này do đó cũng được bao gồm trong sáng chế.

H1 mini-HA A/Texas/36/1991 (SEQ ID NO: 162) được tạo ra từ H1 FL HA A/Texas/36/1991 (SEQ ID NO: 64) bằng cách

1. Loại bỏ vị trí cắt: đột biến R344Q (đánh số đề cập đến trình tự SEQ ID NO: 64)
2. Loại bỏ các gốc S53 đến P321 và đưa cầu nối GGGG vào giữa D52 và K322 (đánh số đề cập đến trình tự SEQ ID NO: 64)
3. Đưa gốc Serin vào vị trí 407, 417 (F407S, L417S; đánh số đề cập đến trình tự SEQ ID NO: 64), Treonin vào vị trí 410 (V410T; đánh số đề cập đến trình tự SEQ ID NO: 64) và gốc Glyxin vào vị trí 414 (F414G; đánh số đề cập đến trình tự SEQ ID NO: 64) trong vòng giữa chuỗi cuộn xoắn A và chuỗi cuộn xoắn CD (các gốc 403-419 trong trình tự SEQ ID: NO: 64)
4. Đưa cầu disulfit vào bằng cách gây đột biến các gốc Arginin 325 và Treonin 437 thành xystein (R325C, T437C; đánh số đề cập đến trình tự SEQ ID NO: 64)
5. Yếu tố làm ổn định bổ sung được đưa vào bằng cách thay thế 419RRMENLNKKVDDGFLD434(đánh số đề cập đến trình tự SEQ ID NO: 64) bằng trình tự RMKQIEDKIEEIESKQ.

H5 mini-HA A/Vietnam/1203/2004 (SEQ ID NO: 163) được tạo ra từ H5 FL HA A/Vietnam/1203/2004 (SEQ ID NO: 158) bằng cách

1. Loại bỏ vị trí cắt. Do H5 FL HA A/Vietnam/1203/2004 (SEQ ID NO: 158) chứa vị trí cắt đa bazơ (341RRRK346) đột biến ở một vị trí là không đủ để ngăn chặn sự cắt protein. Thay vì loại bỏ trình tự 341RRRK345 và đưa vào đột biến R346Q.
2. Loại bỏ các gốc K52 đến P319 và đưa cầu nối GGGG vào giữa K51 và K320 (đánh số đề cập đến trình tự SEQ ID NO: 158)
3. Đưa gốc Serin vào vị trí 409, 419 (F409S, L419S; đánh số đề cập đến trình tự SEQ ID NO: 158), Treonin vào vị trí 412 (V412T; đánh số đề cập đến trình tự SEQ ID NO: 158) và gốc Glyxin vào vị trí 416 (F416G; đánh số đề cập đến trình tự SEQ ID NO: 158) trong vòng giữa chuỗi cuộn xoắn A và chuỗi cuộn xoắn CD (các gốc 405-421 trong trình tự SEQ ID: NO: 158)
4. Đưa cầu disulfit vào bằng cách gây đột biến các gốc Lysin 323 và Treonin 439 thành xystein (K323C, T439C; đánh số đề cập đến trình tự SEQ ID NO: 158)

5. Yếu tố làm ổn định bổ sung được đưa vào bằng cách thay thế 421rrienlnkkmedgfldv437(đánh số đề cập đến trình tự SEQ ID NO: 158) bằng trình tự RMKQIEDKIEEIESKQI.

Các gen mã hóa các trình tự protein của trình tự SEQ ID: NO: 56, 160, 78, 161, 162, 158 và 163 được tổng hợp và tách dòng vào vector biểu hiện pcADN2004 sử dụng các phương pháp thông thường mà người có hiểu biết trung bình trong lĩnh vực đã biết. Vì lý do so sánh trình tự HA chiều dài đủ của H3 A/Hong Kong/1/1968 (SEQ ID NO: 121), cũng như trình tự HA chiều dài đủ của H1 A/Brisbane/59/2007 (SEQ ID NO: 1) có thêm đột biến vị trí cắt R343Q được bao gồm trong thử nghiệm.

Huyền phù tế bào HEK293F (Invitrogen) (10^6 tế bào /ml, 30 ml) được chuyển nhiễm bằng vector biểu hiện (1 μ g/ml) sử dụng 40 μ l 293transfectin làm tác nhân chuyển nhiễm và để cho tiếp tục nhân lên trong 2 ngày. Tế bào được thu gom, chia thành các phần phân ước (0,3 ml, khoảng 3×10^5 tế bào) và các phần phân ước được xử lý bằng huyết thanh đa dòng xuất hiện kháng HA của H1 (Sino Biological Inc. Beijing, China) để thăm dò sự biểu hiện hoặc kháng thể đơn dòng đặc hiệu HA (5 microgram/ml) và kháng thể thứ cấp được sử dụng để nhuộm. Sau đó tế bào được phân tích bằng phương pháp phân loại tế bào liên kết huỳnh quang (FACS) về sự biểu hiện của polypeptit vùng gốc của HA được gắn lên màng theo sáng chế trên bề mặt tế bào. Dải kháng thể đơn dòng có tính đặc hiệu đã biết là liên kết với vùng gốc trong protein chiều dài đủ (CR6261, CR9114) được sử dụng để thăm dò về sự có mặt của epitop bảo thủ và, bằng cách suy luận, sự cuộn gấp chính xác của HA chiều dài đủ và các mini-HA polypeptit theo sáng chế. Kháng thể đơn dòng CR8020 (đã biết là không liên kết với các HA của H1 và H5) và CR9020 (liên kết với vùng đầu của HA từ H1 A/Brisbane/59/2007) cũng được bao gồm trong thử nghiệm này. Các kết quả được biểu hiện dưới dạng tỷ lệ phần trăm tế bào dương tính và Cường độ phát huỳnh quang trung bình (MFI) và được thể hiện trong FIG. 23.

Xử lý tế bào chuyển nhiễm bằng huyết thanh kháng thể đa dòng kháng H1 dẫn đến 20 đến 80% tế bào dương tính về HA chiều dài đủ (thanh liên) và 40-50% tế bào dương tính về các mini-HA. Đối chứng âm FL H3 A/Hong Kong/1/1968 và cM2 chỉ biểu hiện số lượng rất thấp tế bào dương tính. Điều này được phản ánh bằng cường độ phát huỳnh quang trung bình (đáy dải) thể hiện rõ ràng có thể phát hiện được tín hiệu ở tất cả protein HA chiều dài đủ của H1. Tín hiệu về HA chiều dài đủ của H5 vẫn thấp; tuy nhiên, điều này có thể được giải thích là do số lượng tế bào chuyển nhiễm thấp kết hợp với sự giảm nhận diện bởi huyết thanh đa dòng H1. Đối chứng âm FL A/Hong Kong/1/1968 và cM2 thể hiện cường độ ở mức nền.

Cả CR6261 và CR9114, đã biết là các chất liên kết với vùng gốc nhóm 1 mạnh, nhận diện tất cả proteinHA chiều dài đủ và protein mini-HA của Nhóm 1 như được thể hiện bởi số lượng tế bào dương tính lớn (ca. 50 to ca. 95%) và MFI cao. Đây là bằng chứng chứng tỏ rõ ràng rằng các epitop trung hòa của các kháng thể này có mặt trong protein mini-HA, cho thấy rằng cấu trúc ba chiều mà tương đồng mạch với cấu trúc tự nhiên của vùng gốc HA trong HA chiều dài đủ. Như mong đợi, đối chứng âm CR8020 (đặc hiệu đối với HA nhóm 2) không liên kết với HA chiều dài đủ của H1 và H5 hoặc mini-HA của H1 và H5, chứng tỏ rằng liên kết quan sát được của các kháng thể trung hòa CR6261/CR9114 với các protein mini-HA không xuất phát từ tương tác protein-protein đặc hiệu. Quan sát được rõ ràng liên kết giữa HA chiều dài đủ của H3 từ A/Hong Kong/1/1968 và CR9114 hoặc CR8020 từ cả tỷ lệ phần trăm tế bào dương tính và MFI, phù hợp với ghi nhận trước đó và chứng tỏ chức năng của các kháng thể đơn dòng này. Tương tự, đối chứng âm kháng thể CR9020 (yêu tố liên kết với vùng đầu HA đối với A/Brisbane/59/2007) không nhận diện protein mini-HA hoặc protein HA chiều dài đủ, ngoài trừ của HA từ A/Brisbane/59/2007, nhấn mạnh thêm rằng tính đặc hiệu của liên kết quan sát được giữa CR6261 và CR9114.

Kết luận là, bốn polypeptit mới theo sáng chế có nguồn gốc từ HA đã được tạo ra thể hiện chứa các epitop được nhận diện bởi các kháng thể trung hòa CR6261 và CR9114 trong đó không có vùng đầu HA.

Ví dụ 21: Sự bảo vệ kháng lại thử thách gây chết của virus cúm trong chuột bởi polypeptit theo sáng chế

Để xác định xem polypeptit theo sáng chế có khả năng gây đáp ứng miễn dịch bảo vệ chuột không bị chết khi phơi nhiễm với virus cúm mà có thể bị chết, tiến hành thử nghiệm thử thách bằng virus cúm. Các con chuột được gây miễn dịch qua đường tiêm bắp bằng vector biểu hiện mã hóa trình tự SEQ ID NO:78, 161, 45 và 6, cũng như HA chiều dài đủ từ A/Brisbane/59/2007 (SEQ ID NO: 1) chứa đột biến bổ sung R343Q để loại bỏ vị trí cắt. Vector biểu hiện mã hóa cM2 được bao gồm làm đối chứng âm. Tiến hành gây miễn dịch bằng cách sử dụng 50 µg cấu trúc biểu hiện + 50 µg tá được (pUMCV1-GM-CSF) theo quy trình nghiên cứu dưới đây

Quy trình nghiên cứu

Ngày -1 Lấy máu.

Ngày 0 Sử dụng vacxin (tiêm bắp).

Ngày 21 Sử dụng vacxin (tiêm bắp).

Ngày 28 Lấy máu.

Ngày 42 Sử dụng vacxin (tiêm bắp).

Ngày 47 Lấy máu.

Ngày 48 Đo khối lượng, nhiệt độ, điểm lâm sàng và khả năng gây chết hàng ngày.

Ngày 49 Thử thách với liều gây chết của nhiễm virus cúm (qua mũi).

Ngày 49 Chết chủng ngừa còn lại được sử dụng để chuẩn độ sau virus.

Ngày 49-70 Đo khối lượng, nhiệt độ, điểm lâm sàng và khả năng gây chết hàng ngày.

Những con vật có điểm lâm sàng ≥ 3 được kiểm tra hai lần mỗi ngày.

Những con vật có điểm lâm sàng ≥ 4 hoặc nhiệt độ $> 32^{\circ}\text{C}$, bất kể xuất hiện yếu tố nào trước, được loại ngay ra khỏi nghiên cứu này.

Ngày 70 Hy sinh tất cả con chuột.

Nhóm 1-6: Thử thách với PR8 (A/Puerto Rico8/34, H1N1)

Nhóm 1: SEQ ID NO: 78

Nhóm 2: SEQ ID NO: 161

Nhóm 3: SEQ ID NO: 1 R343Q

Nhóm 4: SEQ ID NO: 45

Nhóm 5: SEQ ID NO: 6

Nhóm 6: vector trống

10 con chuột mỗi nhóm. Tổng số 60 con chuột. BALB/c.

Vật liệu và phương pháp:

Chủng virus và nguồn:

Chủng virus cúm PR8 (A/Puerto Rico8/34, H1N1) được lấy từ Virapur (San Diego). Dung dịch gốc 1×10^8 pfu/ml Batch #E2004B.

Điều kiện bảo quản. $-75^{\circ}\text{C} \pm 10^{\circ}\text{C}$. Tủ đông lạnh: -86°C UCT tủ đông lạnh. Dạng nhiệt. Fisher Scientific.

Động vật:

Chuột, BALB/c (không gây bệnh cụ đặc hiệu - Specified Pathogen Free; SPF), con cái. 6-8 tuần tuổi vào Ngày 0 của nghiên cứu ~17-19 gam. Được lấy từ Charles River Laboratories và được xác định bằng 'nhận diện tai' Tất cả các con vật được làm thích nghi và duy trì trong 11 ngày trước khi bắt đầu thử nghiệm.

Sử dụng ADN

Phương pháp hoàn nguyên chất chủng ngừa

Chế phẩm ADN thích hợp, như được liệt kê trên đây được chuẩn bị chia thành các phần phân ước và giữ ở -20°C . Mỗi cấu trúc một phần phân ước được làm tan đá đến nhiệt độ trong phòng ngay trước khi tiêm, rút vào xi lanh và tiêm. Phần còn lại của

mỗi phần phân ước được loại bỏ sau khi hoàn thành tất cả việc tiêm của mỗi vòng gây miễn dịch.

Mức độ liều và phương pháp sử dụng

Các con vật được gây mê bằng cách tiêm vào bụng với 9,75mg Xylasol (Graeub E Dr. AG (www.graeub.com); Cat: 763.02) và 48,75mg Ketazol (Graeub E Dr. AG (www.graeub.com); Cat: 668.51) cho mỗi kg khối lượng cơ thể. 50 µl dung dịch ADN được tiêm sử dụng xi lanh 0,5ml với kim tiêm G29 qua đường trong cơ (i.m.) cơ bốn đầu của mỗi chân sau, tạo ra tổng thể tích là 100 µl được tiêm cho mỗi con chuột. Phần còn lại của mỗi phần phân ước được loại bỏ sau khi hoàn thành tất cả việc tiêm của mỗi vòng gây miễn dịch.

Sử dụng virus:

Phương pháp hoàn nguyên chất chủng ngừa

Vật liệu virus được giữ ở $-75\text{ }^{\circ}\text{C} \pm 10\text{ }^{\circ}\text{C}$ và được làm tan băng trước khi sử dụng. Khi làm tan băng, pha loãng vật liệu trong PBS lạnh ($4\text{ }^{\circ}\text{C}$) tương ứng với 5 LD₅₀/50 µl để thử thách với A/PR/8/34. Virus được pha loãng được giữ trên nước đá cho đến khi sử dụng cho chuột.

Mức độ liều và phương pháp sử dụng

Các con vật được gây mê bằng cách tiêm vào bụng với 9,75mg Xylasol và 48,75mg Ketazol cho mỗi kg khối lượng cơ thể và mỗi con vật nhận 50 µl dung dịch virus qua đường mũi. Vật liệu chưa sử dụng được đưa trở lại phòng thí nghiệm để chuẩn độ sau.

Rút máu và điều chế huyết thanh

Ở các ngày được chỉ định trong quy trình nghiên cứu, trên đây, lấy mẫu máu (lấy mẫu máu trung gian: 100-150 µl thông qua việc đặt ống thông sau obitan, terminal lấy máu qua cardiac puncture: khoảng 300-500 µl). Huyết thanh được phân lập từ máu này bằng cách ly tâm trong 5 phút ở tốc độ 14000 g và giữ ở $-20\text{ }^{\circ}\text{C}$ cho đến khi chuyển lên nước đá khô.

Tính điểm lâm sàng

Các dấu hiệu lâm sàng sau khi thử thách bằng virus được ghi lại bằng hệ thống ghi điểm (1 điểm cho chuột khỏe mạnh; 2 điểm cho chuột biểu hiện các dấu hiệu khó chịu, bao gồm nổi da gà nhẹ, dáng đi hơi thay đổi và tăng vận động bằng chân mót; 3 điểm cho chuột biểu hiện các dấu hiệu nổi da gà mạnh, bụng co thắt, dáng đi thay đổi, giai đoạn không hoạt động, tăng tốc độ thở và đôi khi có tiếng ran (tiếng động lách cách/nổ lách tách); 4 điểm cho chuột có các đặc điểm của nhóm trước được tăng gia tăng, nhưng biểu hiện ít hoạt động và trở nên gần chết; 5 điểm cho chuột chết). Các

con vật được kiểm tra hai ngày một lần với điều kiện chúng nhận được điểm 3. Việc ghi điểm được thực hiện bằng một máy kiểm tra và các con chuột có triệu chứng được biểu hiện không hoàn chỉnh bằng hai điểm số là điểm $\pm 0,5$.

Đo khối lượng cơ thể

Tất cả các con vật được đo khối lượng hàng ngày, bắt đầu vào ngày 48 (số cấp phép 2216). Các con vật cũng được cân trước khi kết thúc nghiên cứu này trong trường hợp chết, tức là khi loại bỏ khỏi nghiên cứu này. Khối lượng cơ thể được ghi lại ở đơn vị gam (g).

Chuẩn độ sau của virus

Liều virus được sử dụng được xác định bằng cách chuẩn độ 8 mẫu lặp lại từ phần chất chủng ngừa còn lại sau hoàn thành việc chủng ngừa cho các con vật. Để chuẩn độ sau của virus sử dụng phép đo TCID₅₀ theo quy trình được đưa ra trong ‘Current Protocols in Immunology, Animal Models of Infectious Disease 19.11.7’.

Các kết quả:

Nghiên cứu được thực hiện không có những khó khăn về kỹ thuật và theo quy trình nghiên cứu đã được xác định. Chuẩn độ sau của các chất chủng ngừa chứa chủng virus cúm PR8 (A/Puerto Rico/8/34, H1N1) thu được IDIC₅₀ sau đây: PR8 (A/Puerto Rico/8/34, H1N1) : $3,2 \times 10^4$ TCID₅₀/ml.

Fig.24A thể hiện đường cong sống sót Kaplan-Meier đối với thử nghiệm này. Việc gây miễn dịch bằng ADN mã hóa polypeptit theo sáng chế có trình tự SEQ ID NO: 45, 6 và 161 dẫn đến sự sống sót 50, 40 và 40% các con chuột được gây nhiễm bằng liều gây chết của virus cúm, tương ứng, chứng tỏ rằng việc gây miễn dịch bằng polypeptit theo sáng chế thực tế là có thể gây ra đáp ứng miễn dịch bảo vệ. Ngược lại các con vật được gây miễn dịch bằng vectơ trống làm đối chứng, tất cả đều ngừng kháng lại sự nhiễm 8 ngày sau khi thử thách bằng virus. Việc gây miễn dịch bằng ADN mã hóa HA chiều dài đủ cùng loài với chủng thử thách (SEQ ID NO: 78) bảo vệ hoàn toàn tất cả các con vật (tức là 100% sống sót) trước thử thách gây chết, trong khi việc gây miễn dịch bằng ADN mã hóa HA chiều dài đủ có nguồn gốc từ chủng khác loài A/Brisbane/59/2007 (SEQ ID NO: 1) chứa thêm đột biến ở vị trí cắt (R343Q; đánh số đề cập đến trình tự SEQ ID NO: 1) dẫn đến 90% các con vật được gây nhiễm sống sót.

Các kết quả thu được từ đường con sống sót cũng phản ánh sự thay đổi khối lượng trung bình và điểm lâm sàng trung bình đối với mỗi nhóm được thể hiện trên Fig.24B và 24C. Các con vật được gây miễn dịch bằng polypeptit theo sáng chế có trình tự SEQ ID NO: 45, 6 và 161 sự giảm khối lượng đến 25-30%, nhưng các con vật

sống sót sau ngày 9 sau khi gây nhiễm thể hiện tăng khối lượng. Đối với các con vật được gây miễn dịch bằng SEQ ID NO: 45 quan sát được điểm lâm sàng giảm từ 4 xuống 3. Các con vật được gây miễn dịch bằng ADN mã hóa HA chiều dài đủ cùng loài với chủng thử thách (SEQ ID NO: 78) không thể hiện sự giảm khối lượng, trong đó các con vật được gây miễn dịch bằng ADN mã hóa HA chiều dài đủ có nguồn gốc từ chủng khác loài A/Brisbane/59/2007 (SEQ ID NO: 1) chứa thêm đột biến ở vị trí cắt (R343Q; đánh số đề cập đến trình tự SEQ ID NO: 1) trải qua sự giảm khối lượng và các triệu chứng lâm sàng nhưng các con sót sót bao gồm toàn bộ. Quan sát được sự giảm khối lượng và các triệu chứng lâm sàng cao nhất đối với nhóm đối chứng được gây miễn dịch bằng vector trống, phù hợp với việc không có con vật nào sống sót trong số các con vật này.

Kết luận là, polypeptit theo sáng chế SEQ ID NO: 6, 45 và 78 có khả năng gây đáp ứng bảo vệ kháng lại thử thách gây chết với H1N1 A/Puerto Rico/8/1934 ở chuột. Lưu ý rằng polypeptit theo sáng chế SEQ ID NO: 6 và 45 có nguồn gốc từ phân tử HA khác loài với chủng thử thách, trong khi SEQ ID NO: 78 có nguồn gốc từ chủng virut cúm cùng loài. Vì vậy polypeptit theo sáng chế có thể gây cảm ứng sự bảo vệ kháng cả việc nhiễm virut cúm cùng loài và khác loài.

Ví dụ 22. *Chọn lọc dải trình tự HA của H1N1 đại diện và thiết kế polypeptit theo sáng chế dựa trên các trình tự này.*

Để thể hiện khả năng ứng dụng rộng rãi của phương pháp được thiết kế được mô tả trong ví dụ 20, phương pháp này được áp dụng đối với dải các trình tự HA0 được chọn lọc mà bao gồm tỷ lệ phần trăm lớn của biến đổi trình tự trong tự nhiên được tìm thấy trong virut H1N1. Việc chọn lọc dải các trình tự HA đại diện từ tập hợp các trình tự HA của H1N1 đã biết ở người, các trình tự trong ví dụ này được đưa vào để chọn lọc số lượng nhỏ nhất các chủng có tính đại diện cao nhất. Để đạt được điều này, tất cả những sự khác nhau giữa các trình tự HA của virut cúm H1N1 ở người có mặt trong dữ liệu về trình tự của virut cúm được định lượng, cấu trúc trong những sự khác nhau này được kiểm tra và xác định các nhóm phụ cùng loài. Từ mỗi nhóm trong số các nhóm này, trình tự đại diện nhất được chọn lọc để đưa vào dải các trình tự này.

Bước đầu tiên trong quy trình này là định lượng sự khác nhau giữa mỗi cặp hoặc các trình tự trong dữ liệu về trình tự được xem xét. PAM250 ngược (rPAM250) nên (Xu, 2004) được sử dụng để định lượng sự khác nhau ở mỗi vị trí axit amin. Sau đó sử dụng phép cộng Oclit để định lượng tổng số sự khác nhau cho một cặp. Tất cả những sự khác nhau của một cặp được sử dụng để tạo nên ma trận đối xứng $n \times n$ của

những sự khác nhau, trong đó tương đương với số lượng chủng virus được nghiên cứu.

Phân tích hiệp nhân tố chính (Principal Coordinates Analysis-PCA) được sử dụng để tạo ma trận về sự khác nhau (Higgins, 1992). PCA dựa trên sự giảm kích thước. Đầu vào ma trận được xem là sự sắp xếp trong không gian n chiều (trong đó n tương đương với số lượng chủng được nghiên cứu). Sau đó độ biến thiên được phân tích và xây dựng theo cách mà đòi hỏi số chiều thấp nhất để bao phủ hầu hết (hoặc toàn bộ) tính biến thiên. Kết quả trong hệ tọa độ m chiều (trong đó m là số chiều để bao phủ hầu hết hoặc toàn bộ độ biến thiên) với hầu hết sự biến đổi trên trục thứ nhất và sau đó giảm. Tất cả trình tự được nghiên cứu được đặt trong hệ tọa độ này. Trong trường hợp mà chỉ cần 2 hoặc 3 chiều, kết quả có thể được vẽ đồ thị trong đồ thị 2D hoặc 3D, tương ứng, trong đó có thể nhìn thấy được sự khác nhau giữa các chủng. Trong trường hợp cần nhiều chiều hơn, có thể được xây dựng đồ thị 3D từ 3 trục đầu tiên, nhưng đồ thị này không bao hàm toàn bộ độ biến thiên, bởi vì một phần biến thiên là trong 4th và các chiều cao hơn.

Các trình tự trong không gian m chiều sau đó được tập hợp thành nhóm, sử dụng cả phương pháp tạo nhóm phân cấp và phương pháp tạo nhóm k trung bình. Sự liên kết trung bình trong các nhóm được sử dụng để thu được các nhóm có độ biến thiên bên trong tương tự, và để tránh tỷ lệ lớn của một nhóm chủng. Việc xếp thành nhóm được thực hiện ở tất cả các mức độ, bắt đầu ở 1 (tất cả các chủng trong một nhóm) cho đến n (mỗi chủng tạo thành nhóm của chính nó). Từ mỗi nhóm, chủng trung tâm nhất được chọn làm đại diện nhất. Sau đó, tập hợp các chủng trung tâm nhất tạo thành dải gồm các chủng đại diện cho nhóm ở mức độ đó. Đối với mỗi mức độ của nhóm, mức độ bao phủ (hoặc tỷ lệ phần trăm của độ biến thiên được giải thích) được ước tính bằng cách tính tổng khoảng cách bình phương của mỗi chủng đến chủng trung tâm của chúng khi so với tổng khoảng cách bình phương của mỗi chủng đến tâm của hệ tọa độ. Sau đó mức độ bao phủ tối thiểu đạt được được thiết lập để là kích thước nhỏ nhất được yêu cầu của dải đại diện.

Ngoài ma trận Xu rPAM250, các giá trị nhỏ được chỉ định đối với sự khác nhau khi một trong hai trình tự có khoảng trống ở một vị trí cụ thể (do các đoạn chèn hoặc đoạn được loại bỏ). Các yếu tố về cân nặng cũng được bao gồm trong quy trình, để tính sự khác nhau lớn nhất trong số các chủng phân lập qua các năm. Độ biến thiên này được xem là độ biến thiên một phần xảy ra thật, một phần bắt nguồn từ các mức độ giám sát/nhận biết khác nhau. Do đó, yếu tố về khối lượng được đặt ở 1 được chia bởi căn bậc hai của số lượng quan sát được trong một năm cụ thể. Yếu tố về khối

lượng được đưa vào xem xét khi tạo cấu trúc không gian 3 chiều, trong quá trình phân tích nhóm và chọn lọc điểm trung tâm, và ước tính mức độ biến thiên được bao phủ.

Trong ví dụ này, các trình tự có cấu trúc được sử dụng, bao gồm các phần của HA mã hóa polypeptit theo sáng chế. Hai tập hợp khác nhau của các trình tự có cấu trúc được tạo ra. Trong tập hợp thứ nhất, các trình tự tự nhiên ngoại trừ trình tự tín hiệu (ví dụ axit amin 1-18), axit amin 53 đến 320 (vùng đầu HA), trình tự xuyên màng (axit amin 530 đến axit amin đầu tận cùng C) (đánh số đề cập đến trình tự SEQ ID NO: 1) hoặc tương đương của các vị trí này trong các trình tự khác được đưa vào xem xét. Trong tập hợp thứ hai axit amin ở vị trí (hoặc vị trí tương đương) 406, 409, 416, 324, 436, 413 không được xem xét, liệu dùng được cải biến theo phương pháp chung được mô tả trong ví dụ 20. Hơn nữa, (tương đương của) các vị trí 419-433 cũng không được tính vào tập hợp thứ hai phản ánh việc bổ sung trình tự làm ổn định dựa trên GCN4 vào trong polypeptit theo sáng chế như được mô tả trong ví dụ 9.

Sử dụng phương pháp được mô tả trên đây 7 trình tự HA được chọn từ tập hợp các trình tự có cấu trúc 1, được chọn lọc, bao phủ 75% độ biến thiên trình tự và 8 trình tự HA từ tập hợp các trình tự có cấu trúc 2 bao phủ 74% độ biến thiên trình tự. Các chủng được liệt kê trong bảng 10. Ba trong số các trình tự được chọn xuất hiện trong cả hai tập hợp, do đó còn lại 13 trình tự HA duy nhất. Các trình tự này được sử dụng để thiết kế polypeptit theo sáng chế theo phương pháp được mô tả trong ví dụ 20. Ngoài ra trình tự làm ổn định GCN4 MKQIEDKIEEIESKQ (SEQ ID NO: 84) được đưa vào ở tương đương của vị trí 419-433 (đánh số đề cập đến trình tự SEQ ID NO: 1), như được mô tả trong ví dụ 9. Polypeptit theo sáng chế được thiết kế dựa trên trình tự HA của H1N1 A/Memphis/20/1978 và H1N1 A/USSR/92/1977 giống nhau, vì là polypeptit theo sáng chế được thiết kế dựa trên trình tự HA của A/Wisconsin/629-D01415/2009 và H1N1 A/Sydney/DD3-55/2010. Do đó trong tổng số duy nhất 10 polypeptit theo sáng chế được thiết kế, và việc ghép và so sánh các trình tự này được thể hiện trên Fig.25.

Các vector biểu hiện chứa ADN mã hóa polypeptit theo sáng chế SEQ ID NO:164 với trình tự SEQ ID NO: 173, cũng như polypeptit theo sáng chế SEQ ID NO: 45 dựa trên các trình tự HA của A/Brisbane/59/2007 và HA chiều dài đủ tương ứng SEQ ID NO: 1 có thêm đột biến vị trí cắt R343Q trong vector biểu hiện pcADN2004 được sử dụng để chuyển nhiễm tế bào HEK293F và tế bào được phân tích bằng FACS như trước. Ngoài kháng thể đơn dòng của người CR6261, CR9114 và CR8020, kháng thể đơn dòng của chuột C179 được biết là trung hòa các chủng cúm A H1 và H2

(Okuna et al., 1993) cũng được bao gồm trong thử nghiệm. Các kết quả được thể hiện trên Fig.26.

Tất cả polypeptit theo sáng chế cũng như trình tự chiều dài đủ của A/Brisbane/59/2007 được biểu hiện trên bề mặt tế bào và nhận diện bởi các kháng thể trung hòa phổ rộng CR6261, CR9114 và C179, nhưng không được nhận diện bởi CR8020. Kháng thể sau được biết là chỉ liên kết với HA từ virus cúm A nhóm 2. Liên kết của kháng thể CR6261, CR9114 và C179 chứng tỏ rằng các epitop trung hòa phổ rộng được bảo toàn tốt trong polypeptit theo sáng chế. Xem xét độ biến thiên trình tự bao hàm trong các trình tự này, đây là bằng chứng rõ ràng về khả năng áp dụng phổ biến của phương pháp do các tác giả sáng chế thiết kế để tạo ra polypeptit theo sáng chế chứa các epitop trung hòa phổ rộng.

Ví dụ 23. Xác định đặc tính của polypeptit theo sáng chế s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145)

Polypeptit tinh sạch theo sáng chế s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) thu được như được mô tả trong ví dụ 13. Để xác nhận sự có mặt của các epitop tạo cấu hình của CR6261 và CR9114 liên kết của các kháng thể này với protein tinh sạch được nghiên cứu bằng phương pháp đo giao thoa lớp sinh học (Octet Red³⁸⁴, Forte Bio). Cuối cùng CR6261, CR9114 và CR8020 được biotinyl hóa được cố định trên bộ cảm biến được phủ streptavidin, bộ cảm biến được cho tiếp xúc đầu tiên với dung dịch polypeptit tinh sạch (250 nM) theo sáng chế để đo tỷ lệ liên kết và sau đó với dung dịch rửa để đo tỷ lệ phân ly. Vì lý do so sánh thử nghiệm này được lặp lại với protein chiều dài đủ (SEQ ID NO 149) cả ở dạng trime và monome của nó. Các kết quả được thể hiện trên Fig.27.

CR6261 được cố định nhận diện cả dạng monome và trime của vùng ngoài của HA chiều dài đủ từ H1N1 A/Brisbane/59/2007 như được chứng minh bằng đáp ứng rõ ràng sau khi tiếp xúc với các protein này trong dung dịch (Fig.27A). Đáp ứng quan sát được đối với protein dạng trime là lớn hơn đáp ứng quan sát được đối với dạng monome (ca. 1,3 nm và 0,9) có cùng trình tự, tác động được gây ra (ít nhất một phần) do kích thước nhỏ của monome so với trime. Liên kết của CR6261 với s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) dẫn đến đáp ứng lớn nhất vào khoảng 0,25 nm. Khi tiếp xúc với dung dịch rửa, quan sát được sự phân ly của phức hợp ở tất cả ba chất phân tích sự giải phóng nhanh nhất quan sát được đối với polypeptit theo sáng chế s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145), và chậm nhất đối với dạng trime của vùng ngoài của HA chiều dài đủ từ H1N1 A/Brisbane/59/2007 (SEQ ID NO 149).

Tương tự với CR9114 được cố định CR6261 cũng nhận diện cả dạng trime và monome của vùng ngoài của HA chiều dài đủ từ H1N1 A/Brisbane/59/2007, cũng như polypeptit theo sáng chế s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145). Đáp ứng là mạnh hơn ở tất cả ba chất phân tích so với CR6261 (1,5, 1,4 và 0,8 nm đối với dạng trime và monome của HA chiều dài đủ (SEQ ID NO: 149) và polypeptit vùng gốc s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145), tương ứng) và khi sự tiếp xúc của phức với đệm rửa, sự giải phóng của kháng nguyên là nhỏ nhất hoặc không thể phát hiện được ở tất cả ba trường hợp. Đối với CR8020 không có đáp ứng nào quan sát được đối với bất kỳ chất phân tích nào, phù hợp với tính đặc hiệu của kháng thể với vùng gốc của virus cúm 2.

Để tiếp tục xác định đặc tính của liên kết của CR6261 và CR9114 với polypeptit vùng gốc tinh sạch, tiến hành chuẩn độ. Cuối cùng CR6261 được cố định chứa bộ cảm biến được cho tiếp xúc với các dung dịch chứa s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) ở nồng độ 500, 250, 125, 63, 31, 16 và 8 nM, tương ứng, và đáp ứng cuối cùng sau 14000s được ghi lại. Các đáp ứng được vẽ đồ thị dưới dạng hàm số của nồng độ polypeptit vùng gốc, và phù hợp với trạng thái ổn định của 1:1 liên kết mẫu được tiến hành, tạo ra sự phân ly ổn định K_d ở ca 190 nM đối với phức hợp CR6261/polypeptit vùng gốc FIG.28A. Tương tự, bộ cảm biến được thay đổi bằng CR9114 cố định được cho tiếp xúc với s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) ở các nồng độ 80, 40, 20, 10, 5, 2,5 và 1,3 nM, tương ứng, và đáp ứng cuối cùng sau 10800s được ghi lại. Làm phù hợp các đáp ứng cuối cùng theo hàm số của nồng độ polypeptit vùng gốc tạo ra trị số K_d bằng 5,4 nM đối với phức CR9114/polypeptit vùng gốc (FIG. 28B).

Kết luận là polypeptit theo sáng chế s-H1-mini2-cụm1+5+6-GCN4 (SEQ ID NO: 145) có khả năng liên kết với kháng thể đơn dòng trung hòa phổ rộng CR6261 và CR9114, xác nhận sự có mặt của các epitop trung hòa tương ứng trong polypeptit vùng gốc này.

Ví dụ 24: *Bảo vệ kháng lại thử thách gây chết bằng virus cúm ở chuột bằng polypeptit theo sáng chế*

Để xác định xem polypeptit theo sáng chế có khả năng gây đáp ứng miễn dịch mà bảo vệ các con chuột trước cái chết khi phơi nhiễm với virus cúm mà nếu không sẽ chết tiến hành thử nghiệm thử thách bằng virus cúm. Các con chuột được gây miễn dịch qua đường tiêm bắp bằng vectơ biểu hiện mã hóa H3 Chiều dài đủ A/Hong Kong/1/1968 (SEQ ID NO: 121), HK68 H3m2-cl9+10+11 (SEQ ID NO: 124) và HK68 H3m2-cl9+10+11+12-GCN4 SEQ ID NO: 130. Tiến hành gây miễn dịch sử

dùng 50 µg expression cấu trúc + 50 µg tá được (pUMCV1-GM-CSF) theo quy trình nghiên cứu dưới đây

Quy trình nghiên cứu

Ngày -1 Lấy máu.

Ngày 0 Sử dụng vacxin (tiêm bắp).

Ngày 21 Sử dụng vacxin (tiêm bắp).

Ngày 28 Lấy máu.

Ngày 42 Sử dụng vacxin (tiêm bắp).

Ngày 47 Lấy máu.

Ngày 48 Đo khối lượng, nhiệt độ, điểm lâm sàng và khả năng gây chết hàng ngày.

Ngày 49 Thử thách với liều gây chết của nhiễm virus cúm (qua mũi).

Ngày 49 Chết chủng ngừa còn lại được sử dụng để chuẩn độ sau virus.

Ngày 49-70 Đo khối lượng, nhiệt độ, điểm lâm sàng và khả năng gây chết hàng ngày.

Những con vật có điểm lâm sàng ≥ 3 được kiểm tra hai lần mỗi ngày.

Những con vật có điểm lâm sàng ≥ 4 hoặc nhiệt độ $\geq 32^\circ \text{C}$, bất kể xuất hiện yếu tố nào trước, được loại ngay ra khỏi nghiên cứu này.

Ngày 70 Hy sinh tất cả con chuột.

Nhóm 7-10: Thử thách với HK68 (A/Hong Kong/1/68, H3N2)

Nhóm 7: SEQ ID NO: 121

Nhóm 8: SEQ ID NO: 130

Nhóm 9: SEQ ID NO: 124

Nhóm 10: vector trống

10 con chuột mỗi nhóm. Tổng số 40 con chuột. BALB/c.

Vật liệu và phương pháp:

Chủng virus và nguồn:

Chủng virus cúm HK68 (A/Hong Kong/1/68) được cung cấp bởi Prof J. Katz (Trung tâm phòng ngừa và kiểm soát bệnh, Atlanta, GA, USA) sau đó được nhân giống bởi Virapur (San Diego). Virus được chuyển qua nhiều lần trong phổi chuột để tăng độc tính ở chuột. Đối chứng phù hợp cho virus này là: Frace et al., Vacxin 1999; 17:2237. Dung dịch gốc 3×10^8 pfu/ml. Lô #F1109A.

Điều kiện bảo quản. $-75^\circ \text{C} \pm 10^\circ \text{C}$. Tủ đông lạnh: -86°C UCT tủ đông lạnh. Dạng nhiệt. Fisher Scientific.

Động vật:

Chuột, BALB/c (Specified Pathogen Free; SPF), con cái. 6-8 tuần tuổi vào Ngày 0 của nghiên cứu ~17-19 gam. Được lấy từ Charles River Laboratories và được xác định bằng ‘nhận diện tai’ Tất cả các con vật được làm thích nghi và duy trì trong 11 ngày trước khi bắt đầu thử nghiệm.

Sử dụng ADN

Phương pháp hoàn nguyên chất chủng ngừa

Chế phẩm ADN thích hợp, như được liệt kê trên đây được chuẩn bị thành các phần phân ước và giữ ở -20 °C. Cho mỗi cấu trúc một phần phân ước được làm tan bằng ở nhiệt độ trong phòng ngay trước khi tiêm, rút vào xi lanh và tiêm. Phần còn lại của mỗi phần phân ước được loại bỏ sau khi hoàn thành tất cả việc tiêm của mỗi vòng gây miễn dịch.

Mức độ liều và phương pháp sử dụng

Các con vật được gây mê bằng cách tiêm vào bụng với 9,75mg Xylasol (Graeub E Dr. AG (www.graeub.com); Cat: 763.02) và 48,75mg Ketazol (Graeub E Dr. AG (www.graeub.com); Cat: 668.51) cho mỗi kg khối lượng cơ thể. 50 µl dung dịch ADN được tiêm sử dụng xi lanh 0,5ml với kim tiêm G29 qua đường trong cơ (i.m.) cơ bốn đầu của mỗi chân sau, tạo ra tổng thể tích là 100 µl được tiêm cho mỗi con chuột. Phần còn lại của mỗi phần phân ước được loại bỏ sau khi hoàn thành tất cả việc tiêm của mỗi vòng gây miễn dịch.

Sử dụng virus:

Phương pháp hoàn nguyên chất chủng ngừa

Vật liệu virus được giữ ở -75 °C ± 10 °C và được làm tan bằng trước khi sử dụng. Khi làm tan bằng, pha loãng vật liệu trong PBS lạnh (4 °C) tương ứng với 10 LD50/50 µl để thử thách với A/HK/1/68. Virus được pha loãng được giữ trên nước đá cho đến khi sử dụng cho chuột.

Mức độ liều và phương pháp sử dụng

Các con vật được gây mê bằng cách tiêm vào bụng với 9,75mg Xylasol và 48,75mg Ketazol cho mỗi kg khối lượng cơ thể và mỗi con vật nhận 50 µl dung dịch virus qua đường mũi. Vật liệu chưa sử dụng được đưa trở lại phòng thí nghiệm để chuẩn độ sau.

Rút máu và điều chế huyết thanh

Ở các ngày được chỉ định trong quy trình nghiên cứu, trên đây, lấy mẫu máu (lấy mẫu máu trung gian: 100-150 µl thông qua việc đặt ống thông sau obitan, lấy máu sau cùng bằng cách chích vào mạch: khoảng 300-500 µl). Huyết thanh được phân lập

từ máu này bằng cách ly tâm trong 5 phút ở tốc độ 14000 g và giữ ở -20 °C cho đến khi chuyển lên nước đá khô.

Tính điểm lâm sàng

Các dấu hiệu lâm sàng sau khi thử thách bằng virus được ghi lại bằng hệ thống ghi điểm (1 điểm cho chuột khỏe mạnh; 2 điểm cho chuột biểu hiện các dấu hiệu khó chịu, bao gồm nổi da gà nhẹ, dáng đi hơi thay đổi và tăng vận động bằng chân mót; 3 điểm cho chuột biểu hiện các dấu hiệu nổi da gà mạnh, bụng co thắt, dáng đi thay đổi, giai đoạn không hoạt động, tăng tốc độ thở và đôi khi có tiếng ran (tiếng động lách cách/nổ lách tách); 4 điểm cho chuột có các đặc điểm của nhóm trước được tăng gia tăng, nhưng biểu hiện ít hoạt động và trở nên gần chết; 5 điểm cho chuột chết). Các con vật được kiểm tra hai ngày một lần với điều kiện chúng nhận được điểm 3. Việc ghi điểm được thực hiện bằng một máy kiểm tra và các con chuột có triệu chứng được biểu hiện không hoàn chỉnh bằng hai điểm số là điểm $\pm 0,5$.

Đo khối lượng cơ thể

Tất cả các con vật được đo khối lượng hàng ngày, bắt đầu vào ngày 48 (số cấp phép 2216). Các con vật cũng được cân trước khi kết thúc nghiên cứu này trong trường hợp chết, tức là khi loại bỏ khỏi nghiên cứu này. Khối lượng cơ thể được ghi lại ở đơn vị gam (g).

Chuẩn độ sau của virus

Liều virus được sử dụng được xác định bằng cách chuẩn độ 8 mẫu lặp lại từ phần chất chủng ngừa còn lại sau hoàn thành việc chủng ngừa cho các con vật. Để chuẩn độ sau của virus sử dụng phép đo TCID₅₀ theo quy trình được đưa ra trong ‘Current Protocols in Immunology, Animal Models of Infectious Disease 19.11.7’.

Các kết quả:

Nghiên cứu được thực hiện không có những khó khăn về kỹ thuật và theo quy trình nghiên cứu đã được xác định. Chuẩn độ sau của các chất chủng ngừa chứa Chủng virus cúm HK68 (A/Hong Kong/1/68, H3N2) thu được IDIC₅₀ sau đây: HK68 (A/Hong Kong/1/68): 1×10^3 TCID₅₀/ml.

Fig.29 thể hiện đáp ứng IgG kháng vùng ngoài của HA từ A/Hong Kong/1/1968 ở ngày 49 như được xác định bằng ELISA. Việc gây miễn dịch bằng ADN mã hóa polypeptit theo sáng chế SEQ ID NO: 124 và SEQ ID NO: 130 gây đáp ứng có thể phát hiện được rõ ràng kháng lại vùng ngoài HA của H3 HK68, trong khi không phát hiện được đáp ứng đối với vectơ trống đối chứng âm. Như mong đợi, quan sát được đáp ứng cao nhất đối với việc gây miễn dịch bằng H3 Chiều dài đủ A/Hong Kong/1/1968 (SEQ ID NO: 121).

Fig.30A thể hiện đường cong sống sót Kaplan-Meier đối với thử nghiệm này. Việc gây miễn dịch bằng ADN mã hóa polypeptit theo sáng chế SEQ ID NO: 124 và 130 dẫn đến sự sống sót 40 và 20% các con chuột được gây nhiễm bằng liều gây chết của virus cúm, tương ứng, chứng tỏ rằng việc gây miễn dịch bằng polypeptit theo sáng chế thực tế là có thể gây ra đáp ứng miễn dịch bảo vệ. Ngược lại các con vật được gây miễn dịch bằng vector trống làm đối chứng tất cả ngừng kháng với sự nhiễm 10 ngày sau khi thử thách bằng virus. Việc gây miễn dịch bằng ADN mã hóa HA chiều dài đủ (SEQ ID NO: 121) bảo vệ hoàn toàn tất cả các con vật (tức là 100% sống sót) trước thử thách gây chết.

Các kết quả thu được từ đường con sống sót cũng phản ánh sự thay đổi khối lượng trung bình và điểm lâm sàng trung bình đối với mỗi nhóm được thể hiện trên Fig.30B và 30C. Các con vật được gây miễn dịch bằng ADN mã hóa polypeptit theo sáng chế SEQ ID NO: 124 và 130 sự giảm khối lượng đến 20%, nhưng các con vật sống sót sau ngày 9 sau khi gây nhiễm thể hiện tăng khối lượng. Các con vật được gây miễn dịch bằng ADN mã hóa HA chiều dài đủ (SEQ ID NO: 121) không thể hiện sự giảm khối lượng. Quan sát được sự giảm khối lượng và triệu chứng lâm sàng cao nhất đối với nhóm đối chứng được gây miễn dịch bằng vector trống, phù hợp với việc không có sự sống sót trong các con vật này.

Kết luận là, polypeptit theo sáng chế SEQ ID NO: 124 và 130 có tính gây miễn dịch và có khả năng gây ra đáp ứng bảo vệ kháng lại thử thách gây chết với H3N2 A/Hong Kong/1/1968 ở chuột.

Ví dụ 25. Thiết kế và xác định đặc điểm của polypeptit vùng gốc dựa trên HA của H3

Để cải thiện polypeptit vùng gốc được mô tả trong ví dụ 12 tập hợp các cấu trúc bổ sung được thiết kế. Hai tập hợp bổ sung chứa đột biến cysteine được thiết kế sẽ cho phép việc hình thành cầu disulfide làm ổn định ở vị trí 53 và 334 (T53C, G334C; cụm 16) và vị trí 39 và 51 (G39C-E51C; cụm 17) (đánh số đề cập đến trình tự SEQ ID NO: 121). Hơn nữa, hai trình tự được chèn vào giữa vị trí 420-421, tức là ở vị trí đầu tận cùng N của chuỗi cuộn xoắn CD- dài (xem Fig.1). Các trình tự chèn này được thiết kế chúng tạo thuận lợi cho sự hình thành cầu disulfide nội-monome giữa các monome riêng trong phân tử trimer. Hai trình tự khác nhau được thiết kế, tức là NATGGCCGG (Cụm 18) và GSGKCCGG (Cụm 19). Trình tự cụm 18 còn bao gồm việc đưa trình tự vị trí glycosyl hóa (tức là NAT) vào trong polypeptit vùng gốc. Trong một số trường hợp vị trí glycosyl hóa cũng được đưa vào ở vị trí 417-419 bằng cách gây đột biến thành NAT.

Sử dụng trình tự của HA chiều dài đủ từ A/Hong Kong/1/1968 làm điểm khởi đầu các sự cải biến được mô tả trên đây được kết hợp với việc loại bỏ cụm S62-P322 để đi đến polypeptit vùng gốc sau:

SEQ ID NO: 174:H3 HK mini2a-cầu nối+cl9 +10+11+12+GCN4T-CG7-1

SEQ ID NO: 175:HK68 H3mini2a-cầu nối+cl9_+10+12+18+GCN4T

SEQ ID NO: 176:HK68 H3mini2a-cầu nối+cl9_+10+12+16+CG7-GCN4T

SEQ ID NO: 177:HK68 H3mini2a-cầu nối+cl9_+10+12+19+GCN4T

SEQ ID NO: 178:HK68 H3mini2a-cầu nối+cl9_+10+12+17+CG7-GCN4T

SEQ ID NO: 179:H3 HK68 mini2a-cầu nối2+cl9_+10+12+GCN4T

Các gen mã hóa các trình tự protein được mô tả trên đây được tổng hợp và tách dòng vào vector biểu hiện pcADN2004 sử dụng các phương pháp thông thường đã biết trong lĩnh vực. Sự biểu hiện trên bề mặt tế bào và liên kết của kháng thể đơn dòng được phân tích bởi phương pháp xếp loại tế bào liên kết huỳnh quang như được mô tả trên đây. Vì lý do so sánh HA chiều dài đủ của H3N2 A/Hong Kong/1/1968 (SEQ ID NO: 121), ngoài việc chứa đột biến R345Q ở vị trí cắt, và SEQ ID NO: 130 (HK68 H3m2-cl9+10+11+12-GCN4) cũng được bao gồm trong thử nghiệm này cũng như đối chứng âm cM2.

Fig.31 thể hiện các kết quả của thử nghiệm này. Tất cả các cấu trúc được biểu hiện trên bề mặt tế bào như được chứng minh bằng các đáp ứng (MFI, dải A; tỷ lệ phần trăm tế bào dương tính, dải B) quan sát được đối với huyết thanh đa dòng kháng H3. Quan sát được liên kết của CR8043 và CR8020 đối với SEQ ID NO: 174, 175, 176, 177, 178, 179, 121 và 130, chứng tỏ rằng đột biến của cụm 16 không góp phần làm ổn định các epitop cấu hình của các kháng thể này. Đối với mAb CR9114 liên kết trên đây nên có thể chỉ quan sát được đối với SEQ ID NO: 174 :H3 HK mini2a-cầu nối+cl9 +10+11+12+GCN4T-CG7-1 và ở mức độ kém hơn với SEQ ID NO: 177 :HK68 H3mini2a-cầu nối+cl9_+10+12+17+CG7-GCN4T. Cả hai trình tự chứa thêm vị trí glycosyl hóa trong vòng B, chứng tỏ rằng hiệu quả làm ổn định của cải biến này lên epitope trung hòa cấu hình của CR9114.

Kết luận là các tác giả sáng chế đã thể hiện rằng bằng phương pháp được mô tả trên đây có thể thu được đối polypeptit vùng gốc theo sáng chế với các typ huyết thanh của nhóm 2, cụ thể là các typ phụ H3. Việc làm ổn định hơn các polypeptit vùng gốc này có thể đạt được bằng đưa vị trí glycosyl hóa vào vòng B. Các trình tự này cũng được bao gồm bởi sáng chế.

Ví dụ 26: Tính gây miễn dịch của polypeptit vùng gốc của HA dựa trên HA của H3

Để đánh giá khả năng sinh miễn dịch của polypeptit vùng gốc, các con chuột được gây miễn dịch bằng vector biểu hiện mã hóa H3 chiều dài đủ từ A/Wisconsin/67/2005 (SEQ ID NO: 89), SEQ ID NO: 105: H3-mini2, SEQ ID NO: 108: H3-mini2-cl9+10+11, SEQ ID NO: 112: H3-mini2-cl9+10+12, SEQ ID NO: 111: H3-mini2-cl9+10+11+12, SEQ ID NO: 114: H3-mini2-cl9+10+11+12-tri, SEQ ID NO: 113: H3-mini2-cl9+10+11+12-GCN4, SEQ ID NO: 119: H3-mini3-cl9+10+11+12+14, SEQ ID NO: 120: H3-mini4-cl9+10+11+12+14. Vector biểu hiện mã hóa cM2 cũng được bao gồm làm đối chứng âm.

Các nhóm gồm 4 con chuột (BALB/c) được gây miễn dịch bằng 50 μ g cấu trúc + 50 μ g tá được (pUMCV1-GM-CSF) tiêm bắp vào ngày 1, 21 và 42. Vào ngày 49 tiến hành lấy máu lần cuối và thu lấy huyết thanh. Đối chứng âm plasmit cM2 được sử dụng bằng súng bắn gen, sử dụng khoảng 10 μ g cấu trúc + khoảng 2 μ g tá được (pUMCV1-GM-CSF) và phác đồ gây miễn dịch tương tự. Huyết thanh được phân tích bằng ELISA sử dụng HA chiều dài đủ tái tổ hợp từ A/Wisconsin/67/2005 và A/Hong Kong/1/1968 (thu được từ Protein Sciences Corporation, Meriden, CT, USA) làm kháng nguyên. Tóm lại, phủ lên đĩa 96 giếng bằng 50 ng HA qua đêm ở 4 °C, sau đó ủ với đệm phong bế (100 μ l PBS, pH 7,4 + 2% sữa không kem) trong 1 giờ ở nhiệt độ trong phòng. Rửa các đĩa bằng PBS + 0,05% Tween-20, và 100 μ l của dãy độ pha loãng 2 lần trong đệm phong bế, bắt đầu từ độ pha loãng 50 lần của huyết thanh được bổ sung. Kháng thể liên kết được phát hiện bằng cách sử dụng IgG của dê kháng chuột liên hợp với HRP, sử dụng các quy trình chuẩn đã được thiết lập trong lĩnh vực này. Các độ chuẩn được so với đường cong chuẩn gồm độ pha loãng theo dãy của kháng thể đơn dòng của chuột liên kết với kháng nguyên HA và được biểu hiện dưới dạng các đơn vị ELISA cho mỗi ml (EU/ml). Các kết quả của ELISA sử dụng HA từ A/Wisconsin/67/2005, A/Hong Kong/1/1968 và A/Perth/16/2009 sau 49 ngày được thể hiện trên Fig.32A, B và C, tương ứng. Huyết thanh thu được từ chuột được gây miễn dịch bằng ADN mã hóa polypeptit vùng gốc được bao gồm trong thử nghiệm này có khả năng nhận diện HA chiều dài đủ từ chủng cùng loài A/Wisconsin/67/2005 và ở mức độ tương tự với HA chiều dài đủ từ chủng khác loài A/Hong Kong/1/1968. Ngược lại, huyết thanh thu được từ các con chuột được gây miễn dịch bằng HA chiều dài đủ từ A/Wisconsin/67/2005 (SEQ ID NO: 89) thể hiện đáp ứng cao hơn với HA của chủng cùng loài so với HA của chủng khác loài từ A/Hong Kong/1/1968 và A/Perth/16/2009.

Kết luận là, dữ liệu này cho thấy rằng polypeptit theo sáng chế có nguồn gốc từ HA của H3 có khả năng gây đáp ứng miễn dịch hướng đến HA chiều dài đủ.

Ví dụ 27: Tính gây miễn dịch của polypeptit vùng gốc của HA dựa trên HA của H3 của A/Hong Kong/1/1968

Để đánh giá khả năng sinh miễn dịch của polypeptit vùng gốc các con chuột được gây miễn dịch bằng vector biểu hiện mã hóa HA chiều dài đủ của H3 từ A/Hong Kong/1/1968 (SEQ ID NO: 121), SEQ ID NO: 124: HK68 H3m2-cl9+10+11, SEQ ID NO: 125: HK68 H3m2-cl9+10+12, SEQ ID NO: 126: HK68 H3m2-cl9+10+11+12, SEQ ID NO: 128: HK68 H3m2-cl9+10+11+12-tri, SEQ ID NO: 130: HK68 H3m2-cl9+10+11+12-GCN4. Vector biểu hiện mã hóa cM2 cũng được bao gồm làm đối chứng âm.

Các nhóm gồm 4 con chuột (BALB/c) được gây miễn dịch bằng 100 µg cấu trúc + 100 µg tá dược (pUMCV1-GM-CSF) tiêm bắp vào ngày 1, 21 và 42. Vào ngày 49 tiến hành lấy máu lần cuối và thu lấy huyết thanh. Đối chứng âm plasmit cM2 được sử dụng bằng súng bắn gen, sử dụng khoảng 10 µg cấu trúc + khoảng 2µg tá dược (pUMCV1-GM-CSF) và phác đồ gây miễn dịch tương tự. Huyết thanh được phân tích bằng ELISA sử dụng HA chiều dài đủ tái tổ hợp từ A/Hong Kong/1/1968 (thu được từ Protein Sciences Corporation, Meriden, CT, USA) làm kháng nguyên. Tóm lại, phủ lên đĩa 96 giếng bằng 50 ng HA qua đêm ở 4 °C, sau đó ủ với đệm phong bế (100 µl PBS, pH 7,4 + 2% sữa không kem) trong 1 giờ ở nhiệt độ trong phòng. Rửa các đĩa bằng PBS + 0,05% Tween-20, và 100 µl của dãy độ pha loãng 2 lần trong đệm phong bế, bắt đầu từ độ pha loãng 50 lần của huyết thanh được bổ sung. Kháng thể liên kết được phát hiện bằng cách sử dụng IgG của dê kháng chuột liên hợp với HRP, sử dụng các quy trình chuẩn đã được thiết lập trong lĩnh vực này. Các độ chuẩn được so với đường cong chuẩn gồm độ pha loãng theo dãy của kháng thể đơn dòng của chuột liên kết với kháng nguyên HA và được biểu hiện dưới dạng các đơn vị ELISA cho mỗi ml (EU/ml).

Các kết quả của ELISA sử dụng HA từ A/Hong Kong/1/1968 sau 49 ngày được thể hiện trên Fig.33. Huyết thanh thu được từ các con chuột được gây miễn dịch bằng ADN mã hóa polypeptit vùng gốc bao gồm trong thử nghiệm này có khả năng nhận diện chủng cùng loài HA chiều dài đủ từ A/Hong Kong/1/1968. Như mong đợi việc gây miễn dịch bằng ADN mã hóa HA chiều dài đủ dẫn đến high kháng thể các độ chuẩn to chủng cùng loài HA protein, trong khi việc gây miễn dịch bằng đối chứng âm là vector biểu hiện mã hóa cM2 không gây cảm ứng kháng thể nhận diện HA chiều dài đủ của chủng A/Hong Kong/1/1968.

Kết luận là, dữ liệu này cho thấy rằng polypeptit theo sáng chế có nguồn gốc từ HA của H3 có khả năng gây đáp ứng miễn dịch hướng đến HA chiều dài đủ của H3.

Ví dụ 28: *Thiết kế polypeptit vùng gốc khác dựa trên HA của H1 có khả năng kích hoạt các kháng thể trung hòa virus cúm nhóm 1 và nhóm 2*

Ví dụ 4 và 6 bộc lộ các polypeptit dựa trên các trình tự H1 mà bộc lộ ổn định epitop của kháng thể CR6261 trung hòa phổ rộng. Với thực tế là CR6261 chỉ trung hòa virus cúm từ nhóm phát sinh chủng loại 1, các polypeptit được thiết kế đối với epitope này có thể không kích hoạt phản ứng mạnh với virus cúm thuộc nhóm phát sinh chủng loại 2. Phương pháp khác để thiết kế các polypeptit theo sáng chế bao gồm các kháng thể trung hòa chéo phổ rộng này là sử dụng trình tự HA của các biến thể của H1 mà giống gần hơn với các trình tự HA của H3 về các đặc tính cấu trúc và sinh hóa của axit amin quan trọng trong epitop. Dựa trên so sánh giữa các cấu trúc của các kháng thể đặc hiệu nhóm và các phân tử (CR6261, F10 và HB36) và kháng thể FI6 của virus cúm gây đại dịch dạng tinh thể (Corti et al, 2011), các tác giả sáng chế đã phát hiện ra rằng đột biến nhóm 1-nhóm 2 T49N (HA2) chỉ có thể được cung cấp bởi FI6 mà không đưa vào va chạm nguyên tử trong không gian. Asparagin ở vị trí 49 của HA2 tồn tại trong hai virus cúm nhóm 1 trong dữ liệu về virus cúm của NCBI: A/swine/Hubei/S1/2009 (ACY06623) và A/swine/Haseluenne/IDT2617/2003 (ABV60697). Do đó, theo một phương án các trình tự H1 làm thành phần cơ sở của sáng chế như được bộc lộ trong ví dụ 4, 6 và 9 là một trong số N49 này chứa trình tự HA. Theo cách khác, trình tự theo sáng chế như được mô tả trong ví dụ 4, 6 và 9 có đột biến bổ sung ở vị trí 49 trong HA2 biến đổi T thành axit amin N. Bảng 7 thể hiện sự ghép và so sánh trình tự của các trình tự HA của H1 minh họa mà có thể được sử dụng làm trình tự khởi đầu đối với polypeptit theo sáng chế.

SEQ ID NO: 180 có nguồn gốc từ SEQ ID NO: 45 bằng cách gây đột biến T392N (đánh số đề cập đến trình tự SEQ ID NO: 1) T392 trong trình tự SEQ ID NO: 1 tương ứng với Treonin ở vị trí 49 trong HA2 như được mô tả trên đây. Gen mã hóa polypeptit theo sáng chế này được tổng hợp và tách dòng vào vector biểu hiện pcADN2004 sử dụng các phương pháp mà người có hiểu biết trung bình trong lĩnh vực đã biết. Sự có mặt của các epitop trung hòa của CR9114 và CR6261 được xác nhận bằng phương pháp phân loại tế bào liên kết huỳnh quang như được mô tả trên đây. Các kết quả được thể hiện trên Fig.34. MFI đối với SEQ ID NO: 180 có thể so sánh được với MFI quan sát được đối với SEQ ID NO: 45 và SEQ ID NO: 1 đối với liên kết với CR6261 và CR9114, trong khi CR9020 (đã biết là có liên kết với vùng đầu của phân tử HA chiều dài đủ) chỉ nhận diện SEQ ID NO: 1, và CR8020 (đặc hiệu đối với HA của virus cúm nhóm 2) không nhận diện SEQ ID NO: 180, SEQ ID NO: 1

hoặc SEQ ID NO: 45. Đối chứng âm cM2 không được nhận diện bởi bất kỳ kháng thể đơn dòng nào được sử dụng trong thử nghiệm này.

Kết luận là, SEQ ID NO: 180, chứa đột biến T392N chứa các epitop trung tính CR6261 và CR9114.

Ví dụ 29. Thiết kế thêm polypeptit theo sáng chế không có trình tự xuyên màng.

HA của virus cúm trong dạng tự nhiên của nó tồn tại dưới dạng trime trên màng tế bào hoặc virus. Theo một số phương án trình tự nội bào và xuyên màng được loại bỏ do đó polypeptit được tiết ra (hòa tan được) được sản sinh sau khi biểu hiện trong tế bào. Phương pháp để biểu hiện và tinh sạch các vùng ngoài của HA được tiết ra đã được mô tả (xem ví dụ Dopheide et al 2009; Ekiert et al 2009, 2011; Stevens et al 2004, 2006; Wilson et al 1981). Người có hiểu biết trung bình trong lĩnh vực sẽ hiểu rằng các phương pháp này cũng có thể được áp dụng trực tiếp đối với polypeptit vùng gốc theo sáng chế để đạt được sự biểu hiện của polypeptit được tiết ra (hòa tan được). Do đó các polypeptit này cũng được bao hàm trong sáng chế. Ví dụ, trong trường hợp polypeptit theo sáng chế có nguồn gốc từ trình tự HA của virus cúm nhóm 1, polypeptit hòa tan được theo sáng chế có thể được tạo ra bằng cách loại bỏ trình tự polypeptit từ gốc (tương đương của) 514 đến đầu tận cùng C (đánh số theo trình tự SEQ ID NO: 1), Theo cách khác, các gốc bổ sung có thể được bao gồm trong polypeptit theo sáng chế, ví dụ bằng cách loại bỏ trình tự từ gốc 515, 516, 517, 518, 519, 520, 521 hoặc 522. Tùy ý, trình tự đuôi his (HHHHHH hoặc HHHHHHH) có thể được thêm vào, cho mục đích tinh sạch, tùy ý được nối thông qua cầu nối. Tùy ý, cầu nối có thể chứa vị trí cắt thủy phân protein để loại bỏ đuôi his sau khi tinh sạch. Polypeptit hòa tan được có thể được làm ổn định hơn nữa bằng cách đưa trình tự đã biết để tạo cấu trúc trime, như trình tự cuộn gấp. Các polypeptit thu được như được mô tả trên đây cũng được bao gồm trong sáng chế.

SEQ ID NO: 181 đến 185 thể hiện các trình tự của polypeptit hòa tan được theo sáng chế có nguồn gốc từ trình tự HA của H1N1 A/Brisbane/59/2007. Tương tự, SEQ ID NO: 186 đến 187 thể hiện các trình tự của polypeptit hòa tan được theo sáng chế trình tự HA của H3N2 A/Hong Kong/1/1968. Người có hiểu biết trung bình trong lĩnh vực sẽ hiểu rằng các trình tự tương đương với polypeptit theo sáng chế có nguồn gốc từ trình tự HA khác của chủng vacxin virus cúm A khác của các typ phụ ví dụ H1, H3, H5 có thể được thiết kế. Cũng sẽ là hiển nhiên đối với người có hiểu biết trung bình trong lĩnh vực rằng 6 histidin ở đầu tận cùng C được gắn vào cho mục đích tinh sạch. Bởi vì các phương pháp tinh sạch khác không sử dụng đuôi his cũng được đề

cập, trình tự 6 histidin là tùy ý, các trình tự không có đuôi tinh sạch này cũng được bao gồm trong sáng chế.

Bảng 1. Các vùng CDR của kháng thể. SEQ ID NO được đưa ra giữa trong ngoặc.

Ab	HC CDR1	HC CDR2	HC CDR3	LC CDR1	LC CDR2	LC CDR3
9114	GGTSNNYA (25)	ISPIFGST (26)	ARHGNYYYSGMDV (27)	DSNIGRRS (28)	SND (29)	AAWDDSLKGAV (30)

Bảng 2. Hoạt tính liên kết chéo của CR9114, như được đo bằng ELISA và FACS. H1= HA tái tổ hợp hòa tan được của H1 A/New Caledonia/20/1999; H3= HA tái tổ hợp hòa tan được của H3 A/Wisconsin/67/2005; H5= HA tái tổ hợp hòa tan được của H5 A/Vietnam/1203/04; H7= HA tái tổ hợp hòa tan được của H7 A/Netherlands/219/2003; H9= HA tái tổ hợp hòa tan được của H9 A/Hong Kong/1073/99; B= HA tái tổ hợp hòa tan được của virus cúm B B/Ohio/01/05; Bệnh dại= glycoprotein virus dại; PER.C6=tế bào PER.C6 không được chuyển nhiễm (đối chứng); mH1= HA của H1 A/New Caledonia/20/1999 được biểu hiện bởi tế bào PER.C6; mH3= HA của H3 A/Wisconsin/67/2005 được biểu hiện bởi tế bào PER.C6; mH7= HA của H7 A/Netherlands/219/2003 được biểu hiện bởi tế bào PER.C6; ND=không được thực hiện. + = liên kết (>10x nền); +/- = liên kết thấp (2-10x nền) - = không có liên kết có thể phát hiện được.

	IgG ELISA							IgG Facs			
	H1	H3	H5	H7	H9	B	Bệnh dại	PerC6	mH1	mH3	mH7
CR9114	+	+	+	+	+	+	-	-	+	+	+
CR4098	-	-	-	-	-	-	+	-	-	-	-

Bảng 3. Hoạt tính trung hòa chéo của CR9114; Các độ chuẩn (được thể hiện ở $\mu\text{g/ml}$) là trị số trung bình nhân của IC50 như được xác định theo phương pháp Spearman-Kärber của ít nhất hai thử nghiệm; >100 = không trung hòa ở nồng độ thử nghiệm cao nhất (100 $\mu\text{g/ml}$).

	Typ phụ	Chủng	CR9114
Nhóm I	H1	A/WSN/33	1,1
		A/New Caledonia/20/99	3,7
		A/Solomon Islands/3/2006	1,8
		A/Brisbane/59/2007	2,6
		A/California/7/2009	0,3
	H2	A/Env/MPU3156/05	8,8

	H5	A/Hong Kong/156/97	0,4
		A/EW/MPF461/07	10,5
	H6	A/EW/MPD411/07	10,5
	H8	A/EW/MPH571/08	8,8
	H9	A/Hong Kong/1073/99	4,4
		A/Ck/HK/SSP176/09	6,3
Nhóm II	H3	A/Hong Kong/1/68	19
		A/Johannesburg/33/94	21,9
		A/Panama/2007/1999	39,9
		A/Hiroshima/52/2005	12,5
		A/Wisconsin/67/2005	32,4
		A/Brisbane/10/2007	5,6
	H4	A/WF/MPA 892/06	0,8
	H7	A/Mallard/Netherlands/12/2000	4,8
		A/New York/107/2003	> 100
	H10	A/Chick/Germany/N/49	15,7
	H14	A/Mallard/Astrakhan/263/1982	> 100

Bảng 4. Liên kết của huyết thanh thu được từ các con chuột được gây miễn dịch bằng HA chiều dài đủ hoặc các cấu trúc mini-HA như được phân tích bởi FACS. Dữ liệu là giá trị trung bình (n=4) về tỷ lệ phần trăm tế bào dương tính sau khi nhuộm, sau đó là cường độ phát huỳnh quang trung bình (trong ngoặc). H1-FL (SEQ ID NO: 1), CL1 (SEQ ID NO: 3), CL1+2 (SEQ ID NO: 4) và CL1+4 (SEQ ID NO: 6). cM2 là đối chứng âm.

gây miễn dịch (huyết thanh) được chuyển nhiễm bằng ADN	cM2	H1-FL	CL1	CL1+2	CL1+4
cM2	47.5 (1817)	4.8 (404)	0.9 (272)	0.7 (263)	0.7 (269)
H1-FL	2.1 (389)	84.1 (7130)	40.0 (1324)	38.7 (1195)	45.5 (1618)
CL1	2.1 (368)	39.1 (1124)	43.9 (1763)	ND	ND
CL1+2	1.7 (348)	47.9 (1616)	ND	51.4 (2472)	ND
CL1+4	1.7 (342)	19.6 (787)	ND	ND	30.0 (1047)

% MFI trung bình nhân

Bảng 5. Axit amin chuẩn, các chữ viết tắt và đặc tính

Axit amin	3-Chữ cái	1-Chữ cái	Tính phân cực của chuỗi bên	Điện tích của chuỗi bên (pH 7,4)
alanin	Ala	A	không phân cực	Trung hòa
arginin	Arg	R	có cực	Điện tích dương
asparagin	Asn	N	có cực	Trung hòa
axit aspartic	Asp	D	có cực	Điện tích âm
cystein	Cys	C	không phân cực	Trung hòa
axit glutamic	Glu	E	có cực	Điện tích âm
glutamin	Gln	Q	có cực	Trung hòa
glyxin	Gly	G	không phân cực	Trung hòa
histidin	His	H	có cực	điện tích dương(10%) trung hòa(90%)
isoleuxin	Ile	I	không phân cực	Trung hòa
leuxin	Leu	L	không phân cực	Trung hòa
lysin	Lys	K	có cực	Điện tích dương
methionine	Met	M	không phân cực	Trung hòa
phenylalanin	Phe	F	không phân cực	Trung hòa
prolin	Pro	P	không phân cực	Trung hòa
serin	Ser	S	có cực	Trung hòa
treonin	Thr	T	có cực	Trung hòa
tryptophan	Trp	W	không phân cực	Trung hòa
tyrosin	Tyr	Y	có cực	Trung hòa
valin	Val	V	không phân cực	Trung hòa

Bảng 6. Trình tự liên ứng đối với H1 402-418 (SEQ ID NO: 17), các biến thể khác trong tự nhiên và đột biến làm ổn định polypeptit theo sáng chế. Một hoặc nhiều đột biến trong trình tự bố mẹ có mặt trong polypeptit theo sáng chế.

vị trí	axit amin	bảo toàn (%)	a.a. khác có trong tự nhiên	đột biến được ưu tiên	Đột biến axit amin phân cực khác	Đột biến axit amin tích điện khác	Đột biến linh động khác
402	M	99,88	T,I				
403	N	99,88	T,D				
404	T	96,24	S,A,M,N,I,V				
405	Q	99,92	H,K				
406	F	99,92	L	S	T, N, Q	R,H,K,D,E	G
407	T	99,73	A,I,K				
408	A	96,71	S,G,V				
409	V	99,45	Q,M,I,G,L	T,Q,G	S, N	R,H,K,D,E	
410	G	99,1	S,N,D,C,V				
411	K	99,8	T,N,R,E				
412	E	99,84	A,G				
413	F	99,92	S,L	S	T,N,Q	R,H,K,D,E	G
414	N	96,52	S,D,I				
415	H/K	56.20/32.76	S,T,N,E,Q,R,D				
416	L	99,8	F,S,I,P	S	T, N, Q	R,H,K,D,E	G
417	E	99,84	A,D,R				
418	K/R	63.95/35.69	S,Q,T,N				

Bảng 7. Sắp xếp và so sánh trình tự của các trình tự H1 theo các phương án cụ thể theo sáng chế

1. A/Solomon Islands/6/2003 (H1N1) (SEQ ID NO: 54)
 2. A/Brisbane/59/2007 (H1N1) (SEQ ID NO: 1)
 3. A/New Caledonia/20/1999(H1N1)(SEQ ID NO: 55)
 4. A/California/07/2009 (H1N1) (SEQ ID NO: 56)
 5. A/swine/Hubei/S1/2009(H1N1) (SEQ ID NO: 57)
 6. A/swine/Haseluenne/IDT2617/2003(H1N1) (SEQ ID NO: 58)
 7. A/NewYork/8/2006(H1N1) (SEQ ID NO: 59)
 8. A/SolomonIslands/3/2006(H1N1) (SEQ ID NO: 60)
 9. A/NewYork/146/2000(H1N1) (SEQ ID NO: 61)
 10. A/NewYork/653/1996(H1N1) (SEQ ID NO: 62)
 11. A/Beijing/262/1995(H1N1)(SEQ ID NO: 63)
 12. A/Texas/36/1991(H1N1) (SEQ ID NO: 64)
 13. A/Singapore/6/1986(H1N1) (SEQ ID NO: 65)
 14. A/Chile/1/1983(H1N1)(SEQ ID NO: 66)
 15. A/Baylor/11515/1982(H1N1)(SEQ ID NO: 67)
 16. A/Brazil/11/1978(H1N1)(SEQ ID NO: 68)
 17. A/USSR/90/1977(H1N1)(SEQ ID NO: 69)
 18. A/NewJersey/8/1976(H1N1)SEQ ID NO: 70)
 19. A/Denver/1957(H1N1)(SEQ ID NO: 71)
 20. A/Albany/4835/1948(H1N1)(SEQ ID NO: 72)
 21. A/FortMonmouth/1/1947(H1N1)(SEQ ID NO: 73)
 22. A/Cameron/1946(H1N1)(SEQ ID NO: 74)
 23. A/Weiss/1943(H1N1)(SEQ ID NO: 75)
 24. A/Iowa/1943(H1N1)(SEQ ID NO: 76)
 25. A/Bellamy/1942(H1N1) (SEQ ID NO: 77)
 26. A/PuertoRico/8/1934(H1N1)(SEQ ID NO: 78)
 27. A/WSN/1933(H1N1)(SEQ ID NO: 79)
 28. A/SouthCarolina/1/1918(H1N1)(SEQ ID NO: 80)
1. MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60

3. MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL ENSHNGKLCL 60
3. MKAKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
4. MKAILVVLLY TFATANADTL CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDKHNGKLCK 60
5. MEAKLFLVLC AFTALKADTF CVGYHANYST HTVD TILEKN VTVTHSVNLL ENSHNGKLCS 60
6. MEAKLFLVLC AFTALKADTI CVGYHANNST DTVD TILEKN VTVTHSVNLL ENNHNGKLCS 60
7. MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
8. MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
9. MKAKLLVLLC AFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
10. MKAKLLVLLC AFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
11. MKAKLLVLLC AFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
12. MKAKLLVLLC AFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
13. MKAKLLVLLC AFTATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
14. MKAKLLVLLC ALSATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDNHNGKLCK 60
15. MKAKLLVLLC ALSATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
16. MKAKLLVLLC ALSATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
17. MKAKLLVLLC ALSATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
18. MKAKLLVLLC AFTATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
19. MKAKLLVLLC ALSATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
20. MKAKLLVLLC ALSATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
21. MKAKLLVLLC ALTATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
22. MKAKLLVLLC ALSATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
23. MKARLLVLLC ALAATDADTI CIGYHANNST DTVD TILEKN VTVTHSVNLL EDSHNGKLCL 60
24. MKARLLVLLC ALAATDADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
25. MKARLLVLLC AIAATDADTI CIGYHANNST DTVD TILEKN VTVTHSVNLL EDSHNGKLCL 60
26. MKANLLVLLC ALAAADADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL 60
27. MKAKLLVLLY AFVATDADTI CIGYHANNST DTVD TIFEKN VAVTHSVNLL EDRHNGKLCK 60
28. MEARLLVLLC AFAATNADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCK 60

※. ※. ※※ . . ※※※. ※※※※※※※※※※ ※※※※※※※. ※※※ ※. ※※※※※※※※※※. ※※※※※※※※※※

1. LKGIAPLQLG NCSVAGWILG NPECELLISR ESWSYIVEKP NPENGTCYPG HFADYEELRE 120
2. LKGIAPLQLG NCSVAGWILG NPECELLISK ESWSYIVEKP NPENGTCYPG HFADYEELRE 120
3. LKGIAPLQLG NCSVAGWILG NPECELLISK ESWSYIVETP NPENGTCYPG YFADYEELRE 120
4. LRGVAPLHLG KCNIAGWILG NPECESLSTA SSWSYIVETP SSDNGTCYPG DFIDYEELRE 120
5. LNGKIPLQLG NCNVAGWILG NPKCDLLTA NSSSYIETS KSKNGACYPG EFADYEELKE 120
6. LNKGAPLQLG NCNVAGWILG NPECDLLTV DSWSYIETS NSKNACYPG EFADYEELRE 120
7. LKGIAPLQLG NCSVAGWILG NPECELLISK ESWSYIVETP NPENGTCYPG YFADYEELRE 120
8. LKGIAPLQLG NCSVAGWILG NPECELLISR ESWSYIVEKP NPENGTCYPG HFADYEELRE 120
9. LKGTAPlQLG NCSIAGWILG NPECESLFSK ESWSYIAETP NPKNGTCYPG YFADYEELRE 120
10. LKGTAPlQLG NCSVAGWILG NPECESLFSK ESWSYIAETP NPENGTCYPG YFADYEELRE 120
11. LKGIAPLQLG NCSVAGWILG NPECESLSK ESWSYIVETP NPENGTCYPG YFADYEELRE 120
12. LKGIAPLQLG NCSVAGWILG NPKCESLFSK ESWSYIAETP NPENGTCYPG YFADYEELRE 120
13. LKGIAPLQLG NCSIAGWILG NPECESLFSK KSWSYIAETP NSENGTCYPG YFADYEELRE 120
14. LKGIAPLQLG KCSIAGWILG NPECESLFSK KSWSYIAETP NSENGTCYPG YFADYEELRE 120
15. LKGIAPLQLG KCSIAGWILG NPECESLFSK KSWSYIAETP NSENGTCYPG YFADYEELRE 120
16. LKGIAPLQLG KCSIAGWILG NPECESLFSK KSWSYIAETP NSENGTCYPG YFADYEELRE 120
17. LKGIAPLQLG KCNIAGWILG NPECESLFSK KSWSYIAETP NSENGTCYPG YFADYEELRE 120
18. LKGIAPLQLG NCSIAGWILG NPECESLFSK KSWSYIAETP NSENGTCYPG YFADYEELRE 120
19. LKGIAPLQLG KCNIAGWILG NPECESLFSK KSWSYIAETP NSENGTCYPG YFADYEELRE 120
20. LKGIAPLQLG KCNIAGWILG NPECESLFSK KSWSYIAETP NSENGTCYPG YFADYEELRE 120
21. LKGIAPLQLG KCNIAGWILG NPECESLFSK RSWSYIAETP NSENGTCYPG DFADYEELRE 120
22. LKGIAPLQLG KCNIAGWILG NPECESLFSK RSWSYIAETP NSENGTCYPG DFADYEELRE 120
23. LKGIAPLQLG KCNIAGWILG NPECESLFSK RSWSYIAETP NSENGTCYPG DFADYEELRE 120
24. LKGIAPLQLG KCNIAGWILG NPECESLFSK RSWSYIAETP NSENGTCYPG DFADYEELRE 120
25. LKGIAPLQLG KCNIAGWILG NPECESLFSK RSWSYIAETP NSENGTCYPG DFADYEELRE 120
26. LKGIAPLQLG KCNIAGWILG NPECESLFSK RSWSYIAETP NSENGTCYPG DFADYEELRE 120
27. LKGIAPLQLG KCNIAGWILG NPECESLFSK RSWSYIAETP NSENGTCYPG DFADYEELRE 120
28. LKGIAPLQLG KCNIAGWILG NPECESLFSK RSWSYIAETP NSENGTCYPG DFADYEELRE 120

[illegible]

1. QLSSVSSFER FEIFPKESSW PNHTTT-GVS ASCSHNGESS FYKNLLWLTG KNGLYPNLSK 179
 2. QLSSVSSFER FEIFPKESSW PNHTVT-GVS ASCSHNGESS FYRNLLWLTG KNGLYPNLSK 179
 3. QLSSVSSFER FEIFPKESSW PNHTVT-GVS ASCSHNGKSS FYRNLLWLTG KNGLYPNLSK 179
 4. QLSSVSSFER FEIFPKTSSW PNHDSNKGVT AACPHAGAKS FYKNLIWLK KGNLYPNLSK 180
 5. QLSTVSSFER FEIFPKAISW PDHDATRGTT VACSHSGVNS FYRNLLSTVK KGNLYPNLSK 180
 6. QLSTVSSFER FEIFPKATSW PNHDTTRGTT ISCSHSGANS FYRNLLWIVK KGNLYPNLSK 180
 7. QLSSVSSFER FEIFPKESSW PNHTVT-GVS ASCSHNGKSS FYRNLLWLTG KNGLYPNLSK 179
 8. QLSSVSSFER FEIFPKESSW PNHTTT-GVS ASCSHNGESS FYKNLLWLTG KNGLYPNLSK 179
 9. QLSSVSSFER FEIFPKDSSW PNHTVTGVT ASCSHNGKSS FYKNLLWLTE KNGLYPNLSK 180
 10. QLSSVSSFER FEIFPKESSW PNHTVTGVT ASCSHNGKSS FYKNLLWLTE KNGLYPNLSK 180
 11. QLSSVSSFER FEIFPKESSW PNHTVT-GVT ASCSHNGKSS FYRNLLWLTE KNGLYPNLSK 179
 12. QLSSVSSFER FEIFPKESSW PNHTVTGVT TSCSHNGKSS FYRNLLWLTG KNGLYPNVSK 180
 13. QLSSVSSFER FEIFPKESSW PNHTVTGVT ASCSHKGRSS FYRNLLWLTG KNGLYPNLSK 180
 14. QLSSVSSFER FEIFPKESSW PKHNVTKGVT AACSHKGRSS FYRNLLWLTE KNGLYPNLSK 180
 15. QLSSVSSFER FEIFPKESSW PKHSVTRGVT ASCSHKGRSS FYRNLLWLTE KNGLYPNLSK 180
 16. QLSSVSSFER FEIFPKERSW PKHNITRGVT ASCSHKGRSS FYRNLLWLTE KNGLYPNLSK 180
 17. QLSSVSSFER FEIFPKERSW PKHNVTRGVT ASCSHKGRSS FYRNLLWLTE KNGLYPNLSK 180
 18. QLSSVSSFER FEIFPKESSW PNHTVTGVT ASCSHKGRSS FYRNLLWLTG KNGLYPNLSK 180
 19. QLSSVSSFER FEIFPKERSW PNHTTR-GVT AACPHARKSS FYKNLLWLTE ANGYPNLSR 179
 20. QLSSVSSFER FEIFPKERSW PKHNITRGVT AACSHKGRSS FYRNLLWLTE KNGLYPNLNK 180
 21. QLSSVSSFER FEIFPKERSW PKHNITRGVT AACSHAGKSS FYKNLLWLTE TDGYPNLSK 180
 22. QLSSVSSFER FEIFPKGRSW PEHNIDIGVT AACSHAGKSS FYKNLLWLTE KDGYPNLNK 180
 23. QLSSVSSFER FEIFPKESSW PKHNTARGVT AACSHAGKSS FYRNLLWLTE KDGYPNLNK 180
 24. QLSSVSSFER FEIFSKESSW PKHTTG-GVT AACSHAGKSS FYRNLLWLTE KDGYPNLNN 179
 25. QLSSVTSFER FEIFPKETSW PKHNTTKGVT AACSHAGKCS FYRNLLWLTE KDGYPNLNN 180
 26. QLSSVSSFER FEIFPKESSW PNHTN-GVT AACSHAGKSS FYRNLLWLTE KEGYPNLSK 179
 27. QLSSVSSLER FEIFPKESSW PNHTFN-GVT VSCSHRGKSS FYRNLLWLTG KEGYPNLSK 179
 28. QLSSVSSFER FEIFPKTSSW PNHETTKGVT AACSYAGASS FYRNLLWLTG KEGYPNLSK 180
- *****:**** *
1. SYANNKEKEV LVLWGVHPP NIGDQRALYH KENAYVSVVS SHYSRKFTPE IAKRPKVRDQ 239
 2. SYANNKEKEV LVLWGVHPP NIGNQKALYH TENAYVSVVS SHYSRKFTPE IAKRPKVRDQ 239
 3. SYVNNKEKEV LVLWGVHPP NIGNQALYH TENAYVSVVS SHYSRRFTPE IAKRPKVRDQ 239
 4. SYINDKKEV LVLWGIHPP TSADQQLYQ NADAYVSVGS SRYSKKFKPE IAKRPKVRX 240
 5. SYTNKKGKEV LVLWGVHPP TDSVQQLYQ NKHTYVSVGS SKYYKRFTPE IAKRPKVRGQ 240
 6. SYTNKKGKEV LVLWGVHPP TDSVQQLYQ NNHTYVSVGS SKYYQRFTPE IAKRPKVRGQ 240
 7. SYANNKEKEV LVLWGVHPP NIGDQRALYH TENAYVSVVS SHYSRRFTPE IAKRPKVRDQ 239
 8. SYANNKEKEV LVLWGVHPP NIGDQRALYH KENAYVSVVS SHYSRKFTPE IAKRPKVRDQ 239
 9. SYVNNKKEV LVLWGVHPP NIGDQRAIYH KENAYVSVVS SHYSRRFTPE IAKRPKVRDQ 240
 10. SYVNNKEKEV LVLWGVHPP NIGDQRAIYH TENAYVSVVS SHYSRRFTPE IAKRPKVRDQ 240
 11. SYVNNKEKEV LVLWGVHPP NIGDQRAIYH TENAYVSVVS SHYSRRFTPE IAKRPKVRDQ 240
 12. SYVNNKEKEV LVLWGVHPP NIGDQRAIYH TENAYVSVVS SHYSRRFTPE IAKRPKVRDQ 240
 13. SYVNNKEKEV LVLWGVHPP NIGDQRAIYH TENAYVSVVS SHYNRRFTPE IAKRPKVRDQ 240
 14. SYVNNKEKEV LVLWGVHPP NIEDQKTIYR KENAYVSVVS SHYNRRFTPE IAKRPKVRDQ 240
 15. SYVNDKEKEV LVLWGVHPP NIEDQKTIYR KENAYVSVVS SHYNRRFTPE IAKRPKVRDQ 240
 16. SYVNNKEKEV LVLWGVHPP NIEDQKTIYR KENAYVSVVS SNYNRRFTPE IAKRPKVRGQ 240
 17. SYVNNKEKEV LVLWGVHPP NIEDQKTIYR KENAYVSVVS SNYNRRFTPE IAKRPKVRGQ 240
 18. SYVNNKEKEV LVLWGVHPP NIGDQRAIYH TENAYVSVVS SHYNRRFTPE IAKRPKVRDQ 240
 19. SYVNNKEKEV LVLWGVHPP NIEQRALYR KENAYVSVVS SNYNRRFTPE IAKRPKVRDQ 239
 20. SYVNNKEKEV LVLWGVHPP NIEDQKTIYR KENAYVSVVS SNYNRRFTPE IAKRPKVRGQ 240
 21. SYVNNKEKEV LVLWGVHPP NIEDQKTIYR KENAYVSVVS SNYNRRFTPE IAKRPKVRGQ 240
 22. SYVNNKEKEV LVLWGVHPP NIENQKTIYR KENAYVSVVS SNYNRRFTPE IAKRPKVRGQ 240
 23. SYVNNKKEV LVLWGVHPP SIKEQQLYQ KENAYVSVVS SNYNRRFTPE IAKRPKVRDQ 240
 24. SYVNNKKEV LVLWGVHPP NIKDQQLYQ KENAYVSVVS SNYNRRFTPE IAKRPKVRGQ 239
 25. SYVNNKKEV LVLWGVHPP NIKDQQLYQ KENAYVSVVS SNYNRRFTPE IAKRPKVRGQ 240
 26. SYVNNKKEV LVLWGIHPP NSKEQQLYQ NENAYVSVVT SNYNRRFTPE IAKRPKVRDQ 239
 27. SYVNNKKEV LVLWGVHPP SSDEQQLYS NGNAYVSVAS SNYNRRFTPE IAKRPKVRDQ 239
 28. SYVNNKKEV LVLWGVHPP TGTDQQLYQ NADAYVSVGS SKYNRRFTPE IAKRPKVRDQ 240
- *****:**** *

```

1. EGRINYYWTL LEPGDTIIFE ANGNIAPRY AFALSRGFGS GIINSNAPMD ECDACKQTPQ 299
2. EGRINYYWTL LEPGDTIIFE ANGNIAPRY AFALSRGFGS GIINSNAPMD KCDACKQTPQ 299
3. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNAPMD ECDACKQTPQ 299
4. EGRMNYWTL VEPGDKITFE ATGNLVVPRY AFAMERNAGS GIISDTPVH DCNTTCQTPK 300
5. AGRMNYWTL FDQGDITIFE ATGNLIAPWH AFALKKGSSS GIMLSAQVH NCTTKCQTPH 300
6. AGRMNYWTL LDQGDITIFE ATGNLIAPWH AFALNKGPSG GIMISDAHVH NCTTKCQTPH 300
7. EGRINYYWTL LEPGDTIIFE ANGNIAPRF AFALSRGFGS GIITSNAPMD ECDACKQTPQ 299
8. EGRINYYWTL LEPGDTIIFE ANGNIAPRY AFALSRGFGS GIINSNAPMD ECDACKQTPQ 299
9. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIISNASMG ECDACKQTPQ 300
10. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNASMG ECDACKQTPQ 300
11. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNAPMN ECDACKQTPQ 299
12. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNASMD ECDACKQTPQ 300
13. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNASMD ECDACKQTPQ 300
14. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNASMD ECDACKQTPQ 300
15. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNVSM ECDACKQTPQ 300
16. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNASMD ECDTKCQTPQ 300
17. AGRINYYWTL LEPGDTIIFE ANGNIAPWH AFALNRGFGS GIITSNASMD ECDTKCQTPQ 300
18. EGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNASMD ECDACKQTPQ 300
19. SGRMNYWTL LEPGDTIIFE ATGNLIAPWY AFALSRGPGS GIITSNAPLD ECDTKCQTPQ 299
20. AGRINYYWTL LEPGDTIIFE ANGNIAPWH AFALSRGFGS GIITSNASMD ECDTKCQTPQ 300
21. AGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALSRDGS GIITSNASMD ECDTKCQTPQ 300
22. AGRINYYWTL LEPGDTIIFE ANGNIAPWY AFALNRGIGS GIITSNASMD ECDTKCQTPQ 300
23. AGRMNYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNASMH ECDTKCQTPQ 300
24. AGRINYYWTL LKPGDTIME ANGNIAPWY AFALSRGFGS GIITSNASMH ECDTKCQTPQ 299
25. AGRMNYWTL LEPGDTIIFE ANGNIAPWY AFALSRGFGS GIITSNASMH ECNTKCQTPQ 300
26. AGRMNYWTL LKPGDTIIFE ANGNIAPMY AFALRRGFGS GIITSNASMH ECNTKCQTPQ 299
27. HGRMNYWTL LEPGDTIIFE ATGNLIAPWY AFALRGFES GIITSNASMH ECNTKCQTPQ 299
28. AGRMNYWTL LEPGDTIIFE ATGNLIAPWY AFALNRGSGS GIITSDAPVH DCNTKCQTPH 300
***** ..*** ** * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

1. GAINSSLPFQ NVHPVTIGEC PKYVRS AKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 359
2. GAINSSLPFQ NVHPVTIGEC PKYVRS AKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 359
3. GAINSSLPFQ NVHPVTIGEC PKYVRS AKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 359
4. GAINTSLPFQ NIHPITIGKC PKYVKSTKLR LATGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
5. GALKNNLPQ NVHLFTIGEC PKYVKSTQLR MATGLRNIPS IQSRGLFGAI AGFIEGGRTG 360
6. GALKNNLPQ NVHPS TIGEC PKYVKSTQLR MATGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
7. GAINSSLPFQ NVHPVTIGEC PKYVRS AKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 359
8. GAINSSLPFQ NVHPVTIGEC PKYVRS AKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 359
9. GAINSSLPFQ NVHPVTIGEC PKYVRSTKLR MVTGLRVNPS IQSRGLFGAI AGFIEGGWTG 360
10. GAINSSLPFQ NVHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
11. GAINSSLPFQ NVHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 359
12. GAINSSLPFQ NVHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
13. GAINSSLPFQ NVHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
14. GAINSSLPFQ NVHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
15. GAINSSLPFQ NVHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
16. GAINSSLPFQ NVHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
17. GAINSSLPFQ NIHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
18. GAINSSLPFQ NVHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
19. GAINSSLPFQ NIHPVTIGEC PKYVRSTKLR MVTGLRNIPS VQSRGLFGAI AGFIEGGWTG 359
20. GAINSSLPFQ NIHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
21. GAINSSLPFQ NIHPVTIGEC PKYVKSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
22. GAINSSLPFQ NIHPFTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWDG 360
23. GAINSSLPFQ NIHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
24. GAINSSLPFQ NIHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 359
25. GAINSSLPFQ NIHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 360
26. GAINSSLPYQ NIHPVTIGEC PKYVRS AKLR MVTGLRNIPS IQSRGLFGAI AGFIEGGWTG 359
27. GSINSNLPQ NIHPVTIGEC PKYVRSTKLR MVTGLRNIPS IQYRGLFGAI AGFIEGGWTG 359
28. GAINSSLPFQ NIHPVTIGEC PKYVRSTKLR MATGLRNIPS IQSRGLFGAI AGFIEGGWTG 360

* * * * *

1. MVDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 2. MVDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 3. MVDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 4. MVDGWYGYHH QNEQGSYAA DLKSTQNAID EITNKVNSVI EKMNTQFTAV GKEFNHLEKR 420
 5. MIDGWYGYHH QNEQGSYAA DQKSTQIAID GINNKANSVI GKMNIQLTSV GKEFNSLEKR 420
 6. MIDGWYGYHH QNEQGSYAA DQKSTQIAID GINNKNVNSII EKMNTQFTSV GKEFNDLEKR 420
 7. MVDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 8. MVDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 9. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSII EKMNTQFTAV GKEFNKLERR 420
 10. MIDGWYGYHH QNEQGSYAA DQKSTQNAID GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 11. MMDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 12. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 13. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 14. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSII EKMNTQFTAV GKEFNKLERR 420
 15. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 16. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 17. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 18. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 19. MMDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 20. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 21. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN WITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 22. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 23. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 24. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 25. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 26. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 27. MIDGWYGYHH QNEQGSYAA DQKSTQNAIN GITNKVNSVI EKMNTQFTAV GKEFNKLERR 419
 28. MIDGWYGYHH QNEQGSYAA DQKSTQNAID GITNKVNSVI EKMNTQFTAV GKEFNKLERR 420
 ..***** ***** * *****. ***** * ***** *****.*.*.*

1. MENLNKKVDD GFIDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 479
 2. MENLNKKVDD GFIDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 479
 3. MENLNKKVDD GFIDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 479
 4. IENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 5. KENLNKTVDD RFLDVWTFNA ELLVLENER TLEFHDNLK SLYEKVKSQ RNNAKEIGNG 480
 6. IENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ RNNAKEIGNG 480
 7. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 479
 8. MENLNKKVDD GFIDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 479
 9. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 10. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 11. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 479
 12. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 13. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 14. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 15. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 16. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 17. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 18. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 19. MENLNKKVDD GFMDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ RNNAKEIGNG 479
 20. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 21. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ RNNAKEIGNG 480
 22. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ RNNAKEIGNG 480
 23. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ RNNAKEIGNG 480
 24. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ RNNAKEIGNG 479
 25. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ RNNAKEIGNG 480
 26. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 479
 27. MENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 479
 28. IENLNKKVDD GFLDIWYNA ELLVLENER TLDFHDSNVK NLYEKVKSQ KNNAKEIGNG 480
 ..***** ***** * *****. ***** * ***** *****.*.*.*

1. CFEFYHKCND ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 539
2. CFEFYHKCND ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 539
3. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 539
4. CFEFYHKCDN TCMESVKNGT YDYPKYSEEA KLNREEIDGV KLESTRIYQI LAIYSTVASS 540
5. CFEFYHKRDN ECLECVKNGT YNYPKYSEES KFNREEIVGV KLESMGIHQI LAIYSTVASS 540
6. CFEFYHKCDN ECMESVKNGT YNYPKYSEES KLNREKIDGV KLESMGVHQI LAIYSTVASS 540
7. CFEFYHKCND ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 539
8. CFEFYHKCND ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 539
9. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
10. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
11. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 539
12. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
13. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
14. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
15. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
16. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
17. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
18. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
19. CFEFYHKCDN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYRI LAIYSTVASS 539
20. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
21. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
22. CFEFYHKCNN ECMESVKNGT YDYPKFSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
23. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
24. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 539
25. CFEFYHKCNN ECMESVKNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 540
26. CFEFYHKCDN ECMESVRNGT YDYPKYSEES KLNREKIDGV KLESMGIYQI LAIYSTVASS 539
27. CFEFYHKCDN ECMESVRNGT YDYPKYSEES KLNREKIDGV KLESMGVYQI LAIYSTVASS 539
28. CFEFYHKCDD ACMESVRNGT YDYPKYSEES KLNREEIDGV KLESMGVYQI LAIYSTVASS 540

*****: ***** *****:**** ***** :*. ***** *

1. LVLLVSLGAI SFWMCSNGSL QCRICI 565
2. LVLLVSLGAI SFWMCSNGSL QCRICI 565
3. LVLLVSLGAI SFWMCSNGSL QCRICI 565
4. LVLLVSLGAI SFWMCSNGSL QCRICI 566
5. LVLLVSLGAI SFWMCSNGSL QCRVCI 566
6. LVLLVSLGAI SFWMCSNGSL QCRICI 566
7. LVLLVSLGAI SFWMCSNGSL QCRICI 565
8. LVLLVSLGAI SFWMCSNGSL QCRICI 565
9. LVLLVSLGAI SFWMCSNGSL QCRICI 566
10. LVLLVSLGAI SFWMCSNGSL QCRICI 566
11. LVLLVSLGAI SFWMCSNGSL QCRICI 565
12. LVLLVSLGAI SFWMCSNGSL QCRICI 566
13. LVLLVSLGAI SFWMCSNGSL QCRICI 566
14. LVLLVSLGAI SFWMCSNGSL QCRICI 566
15. LVLLVSLGAI SFWMCSNGSL QCRICI 566
16. LVLLVSLGAI SFWMCSNGSL QCRICI 566
17. LVLLVSLGAI SFWMCSNGSL QCRICI 566
18. LVLLVSLGAI SFWMCSNGSL QCRICI 566
19. LVLLVSLGAI SFWMCSNGSL QCRICI 565
20. LVLLVSLGAI SFWMCSNGSL QCRICI 566
21. LVLLVSLGAI SFWMCSNGSL QCRICI 566
22. LVLLVSLGAI SFWMCSNGSL QCRICI 566
23. LVLLVSLGAI SFWMCSNGSL QCRICI 566
24. LVLLVSLGAI SFWMCSNGSL QCRICI 565
25. LVLLVSLGAI SFWMCSNGSL QCRICI 566
26. LVLLVSLGAI SFWMCSNGSL QCRICI 565
27. LVLLVSLGAI SFWMCSNGSL QCRICI 565
28. LVLLVSLGAI SFWMCSNGSL QCRICI 566

:** ***** *****

Bảng 9. Trình tự liên ứng đối với H3 401-421 (SEQ ID NO: 104), các biến thể khác trong tự nhiên và đột biến làm ổn định polypeptit theo sáng chế. Một hoặc nhiều đột biến trong trình tự bố mẹ có mặt trong polypeptit theo sáng chế.

Vị trí	axit amin	bảo toàn (%)	Khác có trong tự nhiên	đột biến được ưu tiên	có cực	axit amin tích điện	Linh động
401	I	99,26	V	R		K	
402	E/G	56.2/40.3	K				
403	K	94,85	R				
404	T	99,88	A				
405	N	99,88	S				
406	E	100					
407	K	100					
408	F	100		S	T,N,Q	R,H,K,D,E	G
409	H	100					
410	Q	100					
411	I	100		T	S,N,Q	R,H,K,D,E	G
412	E	100					
413	K	100					
414	E	100					
415	F	100		S	T,N,Q	R,H,K,D,E	G
416	S	99,46	T,L				
417	E	99,71	G,D				
418	V	98,64	I	G	S,T,N,Q	R,H,K,D,E	
419	E	100					
420	G	99,84	E				
421	R	100					

Bảng 10. Khối lượng phân tử theo tính toán và trong thử nghiệm như được xác định từ SEC-MALS. Khối lượng phân tử theo tính toán dựa trên thành phần axit amin của polypeptit theo sáng chế được xử lý, tức là sau khi cắt peptit dẫn đầu. Khối lượng bổ sung do glycosyl hóa không được tính đến.

SEQ ID NO	Khối lượng phân tử theo tính toán (Da)	Khối lượng phân tử trong thử nghiệm (kDa)
144	29137	75
145	29257	50
146	29203	56
147	29119	60
148	29293	66
150	29411	64
151	29283	79

Bảng 11.

Các chủng được chọn	
Tập hợp 1	Tập hợp 2
H1N1 A/Maryland/12/1991	H1N1 A/Maryland/12/1991
H1N1 A/Henry/1936/	H1N1 A/Henry/1936/
H1N1 A/AA/Marton/1943	H1N1 A/AA/Marton/1943
H1N1 A/Memphis/20/1978	H1N1 A/USSR/92/1977
H1N1 A/New York/607/1995	H1N1 A/New York/629/1995
H1N1 A/New Jersey/11/2007	H1N1 A/Virginia/UR06-0549/2007
H1N1 A/Wisconsin/629-D01415/2009	H1N1 A/Texas/UR06-0526/2007
	H1N1 A/Sydney/DD3-55/2010

TÀI LIỆU THAM KHẢO

- Bommakanti et al. (2010), PNAS 107(31): 13701-13706.
- Chen et al. (1995), PNAS 92, 12205-12209.
- Coffman et al. (2010), Immunity 33: 492.
- Corti et al. (2011), Science 333(6044): 850-856.
- Delhaise, et al. (1984), J. Mol. Graph. 2:103–106.
- Devereux et al. (1984), Nucl. Acids Res. 12: 387.
- Dopheide, Ward (1981), J Gen Virol. 367–370. Edgar RC (2004), Các axit nucleic Res. 32, 1792-1797.
- Ekiert et al. (2009), Science 324:246.
- Ekiert et al. (2011), Science 333: 844. Ferguson et al. (2003), Nature 422: 428-443.
- Higgins (1992), Comput Appl Biosci 8:15-22. Lorieau et al. 2010, Proc. Natl. Acad. Sci. USA, 107: 11341.
- Okuno et al. (1993). J Virol. 67:2552-2558. Russel et al. (2004), Virology 325 : 287.
- Samantha et al. (2002) Prot Eng. 15, 659-667.
- Steel et al. (2010), mBio 1(1): 1-9.
- Steven et al. (2004) Science 303: 1866.
- Steven et al. (2006) Science 312: 404. Sun et al., (2010), J. Virol. 84: 8683.
- Suzuki et al. (2005), Prot Eng 11, 1051.
- Throsby et al. (2008), Plos một 12(3): 1-15.
- Wilson et al (1981) Nature 289: 366. Woolfson (2005), Adv Protein Chem 70: 79-112.
- Xu & Miranker (2004) Bioinformatics 8: 1214-1221.

DANH MỤC TRÌNH TỰ

SEQ ID NO 1: H1 chiều dài đủ (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGKLCL LKGIAPLQLG NCSVAGWILG NPECELLISK ESWSYIVEKP 100
 NPENGTCTPG HFADYEELRE QLSSVSSFER FEIFPKESSW PNHTVTGVSA 150
 SCSHNGESSF YRNLLWLTGK NGLYPNLSKS YANNKEKEVL VLWGVHHPN 200
 IGDQKALYHT ENAYVSVVSS HYSRKFTPEI AKRPKVRDQE GRINYYWTLL 250
 EPGDTIIFEA NGNLIAPRYA FALSRGFGSG IINSNAPMDK CDAKCQTPQG 300
 AINSSLPFQN VHPVTIGECP KYVRSACLRLM VTGLRNIPSI QSRGLFGAIA 350
 GFIEGGWTGM VDGWYGYHHQ NEQSGSYAAD QKSTQNAING ITNKVNSVIE 400
 KMNTQFTAVG KEFNKLERRM ENLNKKVDDG FIDIWYNAE LLVLLNERT 450
 LDFHDSNVKN LYEKVKSQK NNAKEIGNGC FEFYHKCNDE CMESVKNGTY 500
 DYPKYSEESK LNREKIDGVK LESMGVYQIL AIYSTVASSL VLLVSLGAIS 550
 FWMCSNGLQ CRICI 565

SEQ ID NO 2: miniHA (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGKLCL **GGG**CDAKCQT PQGAINSSLP FQNVHPVTIG ECPKYVRSK 100
 LRMVTGLRNI PSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQSGY 150
 AADQKSTQNA INGITNKVNS VIEKMNTQFT AVGKEFNKLE RRMENLNKKV 200
 DDGFIDIWY NAELLVLLN ERTLDFHDSN VKNLYEKVKS QLKNNAKEIG 250
 NGCFEFYHKC NDECMESVKN GTYDYPKYSE ESKLNREKID GVKLESMGVY 300
 QILAIYSTVA SSLVLLVSLG AISFWMCSNG SLQCRICI 338

SEQ ID NO 3: miniHA cụm 1 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGK**TCG** **GGG**CDAKCQT PQGAINSSLP FQNVHP**TTT**G ECPKYVRSK 100
 LRMVTGLRNI PSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQSGY 150
 AADQKSTQNA INGITNKVNS VIEKMNTQ**ST** ATGKEFNKSE RRMENLNKKV 200
 DDGFIDIWY NAELLVLLN ERTLDFHDSN VKNLYEKVKS QLKNNAKEIG 250
 NGCFEFYHKC NDECMESVKN GTYDYPKYSE ESKLNREKID GVKLESMGVY 300
 QILAIYSTVA SSLVLLVSLG AISFWMCSNG SLQCRICI 338

SEQ ID NO 4: miniHA cụm 1+2 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGK**TCG** **GGG**CDAKCQT PQGAINSSLP FQNVHP**TTT**G ECP**C**YVRSK 100
 LRMVTGLRNI PSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQSGY 150
 AADQKSTQNA INGITNKVNS VIEKMNT**CST** ATGKEFNKSE RRMENLNKKV 200
 DDGFIDIWY NAELLVLLN ERTLDFHDSN VKNLYEKVKS QLKNNAKEIG 250
 NGCFEFYHKC NDECMESVKN GTYDYPKYSE ESKLNREKID GVKLESMGVY 300
 QILAIYSTVA SSLVLLVSLG AISFWMCSNG SLQCRICI 338

SEQ ID NO 5: miniHA cụm 1+3 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGK**TCG** **GGG**CDAKCQT PQGAINSSLP FQNVHP**TTT**G ECP**C**YVRSK 100
 LRMVTGLRNI PSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQSGY 150
 AADQKSTQNA INGITNKVNS VIEKMNTQ**ST** ATGKECNKSE RRMENLNKKV 200
 DDGFIDIWY NAELLVLLN ERTLDFHDSN VKNLYEKVKS QLKNNAKEIG 250
 NGCFEFYHKC NDECMESVKN GTYDYPKYSE ESKLNREKID GVKLESMGVY 300
 QILAIYSTVA SSLVLLVSLG AISFWMCSNG SLQCRICI 338

SEQ ID NO 6: miniHA cüm1+4 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGKTCG GGGCDAKCQT PQGAINSSLP FQNVHPTTTTG ECPKYVRS AK 100
 LRMVTGLRNI PSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQGSY 150
 AADQKSTQNA INGITNKVNS VIEKMNTQST ATGKEFNKSE RRMENLNKKV 200
 DDGFIDIWYTY NAELLVLEN ERTLDFHDSN VKNLYEKVKS QLKNNAKEIG 250
 NGCFEFYHKC NDECMESVKN GTYDYPKYSE ESKLNREKID GVKLESMGVY 300
 QILAIYSTVA SSLVLLVSLG AISFWMC SNG SLQCRICI 338

SEQ ID NO 7: miniHA cüm1+2+3 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGKTCG GGGCDAKCQT PQGAINSSLP FQNVHPTTTTG ECPCYVRS AK 100
 LRMVTGLRNI PSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQGSY 150
 AADQKSTQNA INGITNKVNS VIEKMNTCST ATGKECNKSE RRMCNLNKKV 200
 DDGFIDIWYTY NAELLVLEN ERTLDFHDSN VKNLYEKVKS QLKNNAKEIG 250
 NGCFEFYHKC NDECMESVKN GTYDYPKYSE ESKLNREKID GVKLESMGVY 300
 QILAIYSTVA SSLVLLVSLG AISFWMC SNG SLQCRICI 338

SEQ ID NO 8: miniHA cüm1+2+3+4 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGKTCG GGGCDAKCQT PQGAINSSLP FQNVHPTTTTG ECPCYVRS AK 100
 LRMVTGLRNI PSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQGSY 150
 AADQKSTQNA INGITNKVNS VIEKMNTCST ATGKECNKSE RRMCNLNKKV 200
 DDGFIDIWYTY NAELLVLEN ERTLDFHDSN VKNLYEKVKS QLKNNAKEIG 250
 NGCFEFYHKC NDECMESVKN GTYDSPKYSE ESKLNREKID GVKLESMGVY 300
 QILAIYSTVA SSLVLLVSLG AISFWMC SNG SLQCRICI 338

SEQ ID NO 9: mini1 cüm11 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGKTCG GGGCDAKCQT PQGAINSSLP FQNVHPTTTTG ECPKYVRS AK 100
 LRMVTGLRNI PSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQGSY 150
 AADQKSTQNA INGITNKVNS VIEKMNTQST ATGKEFNKSE RRIENLNKKI 200
 DDGFIDIWYTY NAELLVLEN ERTLDFHDSN VKNLYEKVKS QLKNNAKEIG 250
 NGCFEFYHKC NDECMESVKN GTYDYPKYSE ESKLNREKID GVKLESMGVY 300
 QILAIYSTVA SSLVLLVSLG AISFWMC SNG SLQCRICI 338

SEQ ID NO 10: mini2 cüm11 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVR SAKLRMVTGL RNIPSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQGSY 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEFN 150
 KSERRIENLN KKIDDGFIDI WTYNAELLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVYQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300

I 301

SEQ ID NO 11: mini3 cüm11 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSGGGGNSS LPFQNVHPTT TGECPKYVRS AKLRMVTGLR NIPSIQSQGL 100
 FGALAGFIEG GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV 150
 NSVIEKMNTQ STATGKEFNK SERRIENLNK KIDDGFIDIW TYNAELLVLL 200
 ENERTLDFHD SNVKNLYEKV KSQLKNNKE IGNGCFEFYH KCNDECMESV 250
 KNGTYDYPKY SEESKLNREK IDGVKLESMG VYQILAIYST VASSLVLLVS 300
 LGAISFWMCS NGSLOCRICI 320

SEQ ID NO 12: mini4 cüm11 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGGGGE CPKYVRS AKL RMVTGLRNIP SIQSQGLFGA IAGFIEGGWT 100
 GMVDGWYGYH HQNEQGS GYA ADQKSTQNAI NGITNKVNSV IEKMNTQSTA 150
TGKEFNKSER RIENLNKKID DGFIDIWTYN AELLVLENE RTLDFHDSNV 200
 KNLVEKVKSQ LKNNKEIGN GCFEFYHKCN DECMESVKNG TYDYPKYSEE 250
 SKLNREKIDG VKLESMGVYQ ILAIYSTVAS SLVLLVSLGA ISFWMCSNGS 300
 LQCRICI 307

SEQ ID NO 13: mini1 cüm11+5 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGKTCG GGGCDAKCQT PQGAINSSLP FQNVHPTTTTG ECPKYVCSAK 100
 LRMVTGLRNI PSIQSQGLFG ALAGFIEGGW TGMVDGWYGY HHQNEQGS GY 150
 AADQKSTQNA INGITNKVNS VIEKMNTQST ATGKEFNKSE RRIENLNKKI 200
 DDGFIDIWCY NAELLVLEN ERTLDFHDSN VKNLYEKVKS QLKNNKEIG 250
 NGCFEFYHKC NDECMESVKNG GTYDYPKYSE ESKLNREKID GVKLESMGVY 300
 QILAIYSTVA SSLVLLVSLG AISFWMCSNG SLOCRICI 338

SEQ ID NO 14: mini2 cüm11+5 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVC SAKLRMVTGL RNIPSIQSQG LFGALAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEFN 150
 KSERRIENLN KKIDDGFIDI WCYNAELLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVYQILAIYS TVASSLVLLV SLGAISFWMC SNGSLOCRIC 300
 I 301

SEQ ID NO 15: mini3 cüm11+5 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSGGGGNSS LPFQNVHPTT TGECPKYVCS AKLRMVTGLR NIPSIQSQGL 100
 FGALAGFIEG GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV 150
 NSVIEKMNTQ STATGKEFNK SERRIENLNK KIDDGFIDIW CYNAELLVLL 200
 ENERTLDFHD SNVKNLYEKV KSQLKNNKE IGNGCFEFYH KCNDECMESV 250
 KNGTYDYPKY SEESKLNREK IDGVKLESMG VYQILAIYST VASSLVLLVS 300
 LGAISFWMCS NGSLOCRICI 320

SEQ ID NO 16: mini4 cụm11+5 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGGGGE CPKYVCSAKL RMVTGLRNIP SIQSQGLFGA IAGFIEGGWT 100
 GMVDGWYGYH HQNEQGSYA ADQKSTQNAI NGITNKVNSV IEKMNTQSTA 150
TGKEFNKSER RIENLNKKID DGFIDIWCYN AELLVLLNE RTLDFHDSNV 200
 KNLYEKVKSQ LKNNAKEIGN GCFEFYHKCN DECMESVKNG TYDYPKYSEE 250
 SKLNREKIDG VKLESMGVYQ ILAIYSTVAS SLVLLVSLGA ISFWMCSNGS 300
 LQCRICI 307

SEQ ID NO 17: Trình tự gốc liên ứng của H1 402-418 (đánh số theo trình tự SEQ ID NO:1)
 402 MNTQFTAVG KEFN(H/K)LE(K/R) 418

>sc09-114 VH PROTEIN (SEQ ID NO: 18)

QVQLVQSGAEVKKPGSSVKVSCKSSGGTSNNYAISWVRQAPGQGLDWMGGISPIFGSTAYA
 QKFQGRVTISADIFSNTAYMELNSLTSEDTAVYFCARHGNYYYYSGMDDVWGQGTITVTVSS

>sc09-114 VL PROTEIN (SEQ ID NO: 19)

SYVLTQPPAVSGTPGQRVTISCSGSDSNIGRRSVNWyQQFPGTAPKLLIYSNDQRPSVVPDRFS
 GSKSGTSASLAISGLQSEDEAEYYCAAWDDSLKGAVFGGGTQLTVL

>CR6261 VH PROTEIN (SEQ ID NO: 20)

EVQLVESGAIEVKKPGSSVKVSCKASGGPFRSYAISWVRQAPGQGP
 EWMGGIIPFGTTKYAPKFQGRVTITADDFAGTVYMESSLRSEDT
 AMYYCAKHMGYQVRETMDVWGKGTTVTVSS

>CR6261 VL PROTEIN (SEQ ID NO: 21)

QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNDYVSWYQQLPGTAPK
 LLIYDNNKRPSGIPDRFSGSKSGTSATLGITGLQTGDEANYCATW
 DRRPTAYVVFGGGTKLTVLG

>sc08-057 VH PROTEIN (SEQ ID NO: 22)

evqlvesggglvqpggslrlscaasgftdsvifmswvrqapkgglecvsiidstyyadsvkgrftisrhnsmtvlfemnsrlpddtavvy
 catesgdfgdqtpyhyamdv

>sc08-057 VL PROTEIN (SEQ ID NO: 23)

qsaltqpasvsgspgqsitistgssgdiggynavswyqhpgkapklmiyevtsrpsgvsdrfsasrgdtasltvsglqaedeahyyccsfa
 dsnili

SEQ ID NO: 24 (STEEL)

MKANLLVLLC ALAAADADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 EDSHNGKLCG GGGCNTKCQT PLGAINSSLP YQNIHPVTIG ECPKYVRS AK 100
 LRMVTGLRNIPSIQSRGLFG ALAGFIEGGW TGMIDGWYGY HHQNEQGSYA 150
 AADQKSTQNA INGITNKVNS VIEKMNIQFT AVGKEFNKLE KRMENLNKKV 200
 DDGFLDIWTY NAELLVLEN ERTLDFHDSN VKNLYEKVKS QLKNNAKEIG 250
 NGCFEFYHKC DNECMESVRN GTYDYPKYSE ESKLNREKVD GVKLESMGIY 300
 QILAIYSTVA SSLVLLVSLG AISFWMCSNG SLQCRICI 338

SEQ ID NO: 31: H7 chiều dài đủ (A/Mallard/Netherlands/12/2000)

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VERTNVPRIC SKGKRTVDLG QCGLLGITG PPQCDQFLEF SADLIERRE 100
 GSDVCYPGKF VNEEALRQIL RESGGIDKET MGFTYSGIRT NGATSACRRS 150
 GSSFYAEMKW LLSNTDNAAF PQMTKSYKNT RKDPALIIWG IHHSGSTTEQ 200
 TKLYGSGNKL ITVGSSNYQQ SFVPSPGARP QVNGQSGRID FHWLILNPND 250
 TVTFSFNGAF IAPDRASFLR GKSMGIQSGV QVDANCEGDC YHSGGTIISN 300
 LPFQININSRA VGKCPRYVKQ ESLLLATGMK NVPEIPKGRG *LFGALAGFIE* 350
NGWEGLIDGW YGFRHQNAQG EGTAADYKST QSAIDQITGK LNRLIEKTNQ 400
QFELIDNEFT EVEKQIGNVI NWTRDSMTEV WSYNAELLVA MENQHTIDLA 450
DSEMKNLYER VKRQLRENAE EDGTGCFEIF HKCDDDCMAS IRNNTYDHSK 500
YREEAMQNR IQIDPVKLSSG YKDVILWFSF GASCFILLAI AMGLVFICVK 550
NGNMRCTICI 560

SEQ ID NO: 32: Trình tự gốc liên ứng của H7 394-414 (đánh số theo trình tự SEQ ID NO: 31)

394 LI(E/D/G)KTNQQFELIDNEF (N/T/S) E (I/V) E (Q/K) 414

SEQ ID NO: 33: H7-mini2 (A/Mallard/Netherlands/12/2000)

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VEGGGGGRYVK QESLLLATGM KNPPEIPKGQ *GLFGALAGFI ENGWEGLIDG* 100
WYGFRHQNAQ GEGTAADYKS TQSAIDQITG KLNRLIEKTN QQFELIDNEF 150
TEVEKQIGNV INWTRDSMTE VWSYNAELLV AMENQHTIDL ADSEMKNLYE 200
RVKRQLRENA EEDGTGCFEI FHKCDDDCMA SIRNNTYDHS KYREEAMQNR 250
IQIDPVKLSS GYKDVILWFS FGASCFILLA IAMGLVFICV KNGNMRCTIC 300
I 301

SEQ ID NO: 34: H7-mini2-cum15

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VEGGGGGRYVK QESLLLATGM KNPPEIPKGQ *GLFGALAGFI ENGWEGLIDG* 100
WYGFRHQNAQ GEGTAADYKS TQSAIDQITG KLNRLIEKTN QQSELTDNES 150
TEVEKQIGNV INWTRDSMTE VWSYNAELLV AMENQHTIDL ADSEMKNLYE 200
RVKRQLRENA EEDGTGCFEI FHKCDDDCMA SIRNNTYDHS KYREEAMQNR 250
IQIDPVKLSS GYKDVILWFS FGASCFILLA IAMGLVFICV KNGNMRCTIC 300
I 301

SEQ ID NO: 35: H7-mini2-cum15+16

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VEGGGGGRYVK QESLLLATGM KNPPEIPKGQ *GLFGALAGFI ENGWEGLIDG* 100
WYGFRHQNAQ GEGTAADYKS TQSAIDQITG KLNRLIEKTN QQSENTDNES 150
TENEKQIGNV INWTRDSMTE VWSYNAELLV AMENQHTIDL ADSEMKNLYE 200
RVKRQLRENA EEDGTGCFEI FHKCDDDCMA SIRNNTYDHS KYREEAMQNR 250
IQIDPVKLSS GYKDVILWFS FGASCFILLA IAMGLVFICV KNGNMRCTIC 300
I 301

SEQ ID NO: 36: H7-mini2-cum17

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VEGGGGGRYVC QESLLLATGM KNVPEIPKGQ GLFGALAGFI ENGWEGLIDG 100
 WYGFRHQNAQ GEGTAADYKS TQSAIDQITG KLNRLIEKTN QQFELIDNEF 150
 TEVEKQIGNV INWTRDSMTE VWCYNAELLV AMENQHTIDL ADSEMKNKLYE 200
 RVKRQLRENA EEDGTGCFEI FHKCDDDCMA SIRNNTYDHS KYREEAMQNR 250
 IQIDPVKLSS GYKDVILWFS FGASCFILLA IAMGLVFICV KNGNMRCTIC 300
 I 301

SEQ ID NO: 37: H7-mini2-cum15+16+17

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VEGGGGGRYVC QESLLLATGM KNVPEIPKGQ GLFGALAGFI ENGWEGLIDG 100
 WYGFRHQNAQ GEGTAADYKS TQSAIDQITG KLNRLIEKTN QOSENTDNES 150
 TENEKQIGNV INWTRDSMTE VWCYNAELLV AMENQHTIDL ADSEMKNKLYE 200
 RVKRQLRENA EEDGTGCFEI FHKCDDDCMA SIRNNTYDHS KYREEAMQNR 250
 IQIDPVKLSS GYKDVILWFS FGASCFILLA IAMGLVFICV KNGNMRCTIC 300
 I 301

SEQ ID NO: 38: H7-mini2-cum15+16+17+18

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VEGGGGGRYVC QESLLLATGM KNVPEIPKGQ GLFGALAGFI ENGWEGLIDG 100
 WYGFRHQNAQ GEGTAADYKS TQSAIDQITG KLNRLIEKTN QOSENTDNES 150
 TENEKQIGNI INWIRDIMTE IWCYNAELLV AMENQHTIDL ADSEMKNKLYE 200
 RVKRQLRENA EEDGTGCFEI FHKCDDDCMA SIRNNTYDHS KYREEAMQNR 250
 IQIDPVKLSS GYKDVILWFS FGASCFILLA IAMGLVFICV KNGNMRCTIC 300
 I 301

SEQ ID NO: 39: H7-mini2-cum15+16+17+tri

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VEGGGGGRYVC QESLLLATGM KNVPEIPKGQ GLFGALAGFI ENGWEGLIDG 100
 WYGFRHQNAQ GEGTAADYKS TQSAIDQITG KLNRLIEKTN QOSENTDNES 150
 TENEKQIEAI EKKIEAIMTE IWCYNAELLV AMENQHTIDL ADSEMKNKLYE 200
 RVKRQLRENA EEDGTGCFEI FHKCDDDCMA SIRNNTYDHS KYREEAMQNR 250
 IQIDPVKLSS GYKDVILWFS FGASCFILLA IAMGLVFICV KNGNMRCTIC 300
 I 301

SEQ ID NO: 40: H7-mini5

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VERGGGGRY VKQESLLLAT GMKNVPEIPK GQGLFGALAG FIENGWEGLI 100
 DGWYGFRHQNAQ GEGTAADY KSTQSAIDQI TGKLNRLIEK TNQQFELIDN 150
 EFTEVEKQIG NVINWTRDSM TEVWSYNAEL LVAMENQHTI DLADSEMKNL 200
 YERVKRQLRE NAEEDGTGCF EIFHKCDDDC MASIRNNTYD HSKYREEAMQ 250
 NRIQIDPVKL SSGYKDVILW FSFGASCFIL LAIAMGLVFI CVKNGNMRCT 300
 ICI 303

SEQ ID NO: 41: H7-mini5-cum15+16

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VERGGGGPRY VKQESLLLAT GMKNVPEIPK GQGLFGALAG FIENGWEGLI 100
 DGWYGFRHQ N AQEGTAADY KSTQSAIDQI TGKLNRLIEK TNOQSENTDN 150
 ESTENEKQIG NVINWTRDSM TEVWSYNAEL LVAMENQHTI DLADSEMKNL 200
 YERVKRQLRE NAEEDGTGCF EIFHKCDDDC MASIRNNTYD HSKYREEAMQ 250
 NRIQIDPVKL SSGYKDVILW FSFGASC FIL LAIAMGLVFI CVKNGNMRCT 300
 ICI 303

SEQ ID NO: 42: H7-mini5-cum17

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VERGGGGPRY CKQESLLLAT GMKNVPEIPK GQGLFGALAG FIENGWEGLI 100
 DGWYGFRHQ N AQEGTAADY KSTQSAIDQI TGKLNRLIEK TNOQFELIDN 150
 EFTEVEKQIG NVINWTRDSM TEVWCYNAEL LVAMENQHTI DLADSEMKNL 200
 YERVKRQLRE NAEEDGTGCF EIFHKCDDDC MASIRNNTYD HSKYREEAMQ 250
 NRIQIDPVKL SSGYKDVILW FSFGASC FIL LAIAMGLVFI CVKNGNMRCT 300
 ICI 303

SEQ ID NO: 43: H7-mini5-cum15+16+17+18

MNTQILVFAL MAIPTNADK ICLGHHAVSN GTKVNTLTER GVEVVNATET 50
 VERGGGGPRY VCQESLLLAT GMKNVPEIPK GQGLFGALAG FIENGWEGLI 100
 DGWYGFRHQ N AQEGTAADY KSTQSAIDQI TGKLNRLIEK TNOQSENTDN 150
 ESTENEKQIG NINWIRDIM TEIWCYNAEL LVAMENQHTI DLADSEMKNL 200
 YERVKRQLRE NAEEDGTGCF EIFHKCDDDC MASIRNNTYD HSKYREEAMQ 250
 NRIQIDPVKL SSGYKDVILW FSFGASC FIL LAIAMGLVFI CVKNGNMRCT 300
 ICI 303

SEQ ID NO: 44: H1-mini2-cum1+5+6-trim

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGKYVC SAKLRMVTGL RNPSIQSQG LFGALAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
 KSEIEAIEKK IEAIEKKIEI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKNDECME S VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVYQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 45: H1-mini2-cum1+5+6-GCN4

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGKYVC SAKLRMVTGL RNPSIQSQG LFGALAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
 KSERMKOIED KIEEIESKOI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKNDECME S VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVYQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 46: mini2-cum1+5+6 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVC SAKLRMTGL RNIPSIQSQG LFGALAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
 KSERRMENLN KKVDDGFIDI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 47: mini2-cum1+5+6 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVC SAKLRMTGL RNIPSIQSQG LFGALAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
 KSERRIENLN KKVDDGFIDI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 48: mini2-cum1+5 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVC SAKLRMTGL RNIPSIQSQG LFGALAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEFN 150
 KSERRMENLN KKVDDGFIDI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 49: H1-mini2-cum1+5+6-trim2

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVC SAKLRMTGL RNIPSIQSQG LFGALAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
 KSERIEAIEK KIEAIEKKI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 50: H1-mini2-cum1+5+6-trim3

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVC SAKLRMTGL RNIPSIQSQG LFGALAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
 KSERIEAIE KIEAIEKKI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 51: H1-mini2-cum1+5+6-GCN4t2

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVC SAKLRMVTGL RNIPSIQSQG LFGAIAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
KSERRMKOIE DKIEEIESKI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVYQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 52: H1-mini2-cum1+5+6-GCN4t3

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVC SAKLRMVTGL RNIPSIQSQG LFGAIAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
KSRMKOIEDK IEEIESKOKI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVYQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 53: H1-mini2-cum1+5+6-IleTri

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGGKYVC SAKLRMVTGL RNIPSIQSQG LFGAIAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
KSERRIENIN KKIDDIFIDI WCYNALLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVYQILAIYS TVASSLVLLV SLGAISFWMC SNGSLQCRIC 300
 I 301

SEQ ID NO: 89: H3 chiều dài đủ A/Wisconsin/67/2005

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGICDS PHQILDGENC TLIDALLGDP QCDGFQNKKW 100
 DLFVERSKAY SNCYPYDVPD YASRLSLVAS SGTLEFNDES FNWTGVTQNG 150
 TSSSCKRRSN NSFFSRLNWL THLKFKYPAL NVTMPNNEKF DKLYIWGVHH 200
 PVTDNNDQIFL YAQASGRITV STKRSQQTVI PNIGSRPRIR NIPSRISIW 250
 TIVKPGDILL INSTGNLIAP RGYFKIRSGK SSIMRSDAPI GKCENSECITP 300
 NGSIPNDKPF QNVNRITYGA CORYVKQNTL KLATGMRNVP EKQTRGIFGA 350
 IAGFIENGWE GMVDGWYGFR HQNSEGIGQA ADLKSTQAAI NQINGKLNRL 400
 IGKTNEKFHQ IEKEFSEVEG RIQDLEKYVE DTKIDLWSYN AELLVALENQ 450
 HTIDLTDSEM NKLFFERTKKQ LRENAEDMGN GCFKIYHKCD NACIGSIRNG 500
 TYDHDVYRDE ALNNRFQIKG VELKSGYKDW ILWISFAISC FLLCVLLGF 550
 IMWACQKGN RCNICI 566

SEQ ID NO: 90: mini-H3

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGICGG GGCNSECITP NGSIPNDKPF QNVNRITYGA 100
 CORYVKQNTL KLATGMRNVP EKQTQGIFGA IAGFIENGWE GMVDGWYGFR 150
 HQNSEGIGQA ADLKSTQAAI NQINGKLNRL IGKTNEKFHQ IEKEFSEVEG 200
 RIQDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM NKLFFERTKKQ 250
 LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
 VELKSGYKDW ILWISFAISC FLLCVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 91: mini-H3 cum1

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGTCGG GGCNSECITP NGSIPNDKPF QNVNRITYGA 100
 CORYVKQNTL KLATGMRNVP EKQTQGIFGA IAGFIENGWE GMVDGWYGFR 150
 HQNSEGIGQA ADLKSTQAAI NQINGKLNRL IGKTNEKTSQ IEKEFSESEG 200

RIQDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
 LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
 VELKSGYKDW ILWISFAISC FLLCVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 92: mini-H3 cüm1+2

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGC 100
 CPRYVKQNTL KLATGMRNVP EKQTQGIFGA LAGFIENGWE GMVDGWYGFR 150
 HQNSEGIGQA ADLKSTQAAI NQINGKLNRL IGKTNCKTSQ IEKEFSESEG 200

RIQDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
 LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
 VELKSGYKDW ILWISFAISC FLLCVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 93: mini-H3 cüm1+3

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGA 100
 CPRYVCQNTL KLATGMRNVP EKQTQGIFGA LAGFIENGWE GMVDGWYGFR 150
 HQNSEGIGQA ADLKSTQAAI NQINGKLNRL IGKTNEKTSQ IEKEFSESEG 200

RIQDLEKYVE DTKIALWCYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
 LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
 VELKSGYKDW ILWISFAISC FLLCVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 94: mini-H3 cüm1+4

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGA 100
 CPRYVKQNTL KLATGMRNVP EKQTQGIFGA LAGFIENGWE GMVDGWYGFR 150
 HQNSEGIGQA ADLKSTQAAI NQINGKKNRL TGKTNEKTSQ IEKEFSESEG 200

RIQDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
 LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
 VELKSGYKDW ILWISFAISC FLLCVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 95: mini-H3 cüm1+5 N60A

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGA 100
 CPRYVKQNTL KLATGMRNVP EKQTQGIFGA LAGFIENGWE GMVDGWYGFR 150
 HQNSEGIGQA ADLKSTQAAI NQINGKLNRL IGKTAEKTSQ IEKEFSESEG 200

RIQDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
 LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
 VELKSGYKDW ILWISFAISC FLLCVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 96: mini-H3 cüm1+5 N60D

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGA 100
 CPRYVKQNTL KLATGMRNVP EKQTQGIFGA LAGFIENGWE GMVDGWYGFR 150
 HQNSEGIGQA ADLKSTQAAI NQINGKLNRL IGKTDEKTSQ IEKEFSESEG 200

RIQDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
 LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
 VELKSGYKDW ILWISFAISC FLLCVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 97: mini-H3 cüm1+5 N60E

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGA 100
 CPRYVKQNTL KLATGMRNVP EKQTQGIFGA LAGFIENGWE GMVDGWYGFR 150
 HQNSEGIGQA ADLKSTQAAI NQINGKLNRL IGKTEEKTSQ IEKEFSESEG 200

RIQDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250

LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
VELKSGYKDW ILWISFAISC FLLCVVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 98: mini-H3 cüm1+6

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGA 100
CPRYVKQNTL KLATGMRNVP EKQTQGIFGA IAGFIENGWE GMVDGWYGFR 150
HQNSEGIGQA ADLKSTQAAI NQINGKLNRL IGKTNEKTSQ IEKECSESEG 200
RICDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
VELKSGYKDW ILWISFAISC FLLCVVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 99: mini-H3 cüm1+7

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGA 100
CPRYVKQNTL KLATGMRNVP EKQTQGIFGA IAGFIENGWE GMVDGWYGFR 150
HQNSEGIGQA ADLKSTQAAI NQINGKLNRL IGKTNEKTSQ IEKEFSESEG 200
RIQDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
LRENAEDMGN GCFKIYHKCD NACIESIRNG TYDHDVYRDE ALNNRFQIKG 300
VELKSGYKDW ILWISFAISC FLLCVVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 100: mini-H3 cüm1+2+3+4+5+6+7-N405E

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGC 100
CPRYVCQNTL KLATGMRNVP EKQTQGIFGA IAGFIENGWE GMVDGWYGFR 150
HQNSEGIGQA ADLKSTQAAI NQINGKKNRL TGKTECKTSQ IEKECSESEG 200
RICDLEKYVE DTKIALWCYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
LRENAEDMGN GCFKIYHKCD NACIESIRNG TYDHDVYRDE ALNNRFQIKG 300
VELKSGYKDW ILWISFAISC FLLCVVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 101: mini-H3 cüm1+2+3+4+5+6+7-N405A

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGC 100
CPRYVCQNTL KLATGMRNVP EKQTQGIFGA IAGFIENGWE GMVDGWYGFR 150
HQNSEGIGQA ADLKSTQAAI NQINGKKNRL TGKTECKTSQ IEKECSESEG 200
RICDLEKYVE DTKIALWCYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
LRENAEDMGN GCFKIYHKCD NACIESIRNG TYDHDVYRDE ALNNRFQIKG 300
VELKSGYKDW ILWISFAISC FLLCVVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 102: mini-H3 cüm1+2+3+4+5+6+7-N405D

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGC 100
CPRYVCQNTL KLATGMRNVP EKQTQGIFGA IAGFIENGWE GMVDGWYGFR 150
HQNSEGIGQA ADLKSTQAAI NQINGKKNRL TGKTDCKTSQ IEKECSESEG 200
RICDLEKYVE DTKIALWCYN AELLVALENQ HTIDLTDSEM NKL FERTKKQ 250
LRENAEDMGN GCFKIYHKCD NACIESIRNG TYDHDVYRDE ALNNRFQIKG 300
VELKSGYKDW ILWISFAISC FLLCVVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 103: mini-H3 cüm1+8

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
EVTNATELVQ SSSTGGTCGG GGCNSECTTP NGSIPNDKPF QNVNRQTYGA 100
CPRYVKQNTL KLATGMRNVP EKQTQGICGA IAGFIENGWE GMVDGWYGFR 150
HQNSEGIGQA ADLKCTQAAI NQINGKLNRL IGKTNEKTSQ IEKEFSESEG 200
RIQDLEKYVE DTKIDLWSYN AELLVALENQ HTIDLTDSEM CKCFERTKKQ 250
LRENAEDMGN GCFKIYHKCD NACIGSIRNG TYDHDVYRDE ALNNRFQIKG 300
VELKSGYKDW ILWISFAISC FLLCVVLLGF IMWACQKGN RCNICI 346

SEQ ID NO: 104: Trình tự gốc liên ứng của H3 401-421 (đánh số theo trình tự SEQ ID NO: 1)

401 I(E/G)KTNEKFHQIEKEFSEVEGR 421

SEQ ID NO: 105: H3-mini2

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHA VPNGT IVKTITNDQI 50
 EVTNATELVQ **SGGGG**RYVKQ NTLKLATGMR NVPEKQT**QGI** FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRLIGKTNEK 150
 FHQIEKEFSE VEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300
 GNIRCNICI 309

SEQ ID NO: 106: H3-mini2-cl9+10

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHA VPNGT IVKTITNDQI 50
 EVTNATELVQ **SGGGG**RYVKQ NTLKLATGMR NVPEKQT**QGI** FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRLIGKTNEK 150
SHQTEKESSE GEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300
 GNIRCNICI 309

SEQ ID NO: 107: H3-mini2-cl9+11

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHA VPNGT IVKTITNDQI 50
 EVTNATELVQ **SGGGG**RYVKQ NTLKLATGMR NVPEKQT**QGI** FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRL**R**GKTNEK 150
SHQTEKESSE VEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300
 GNIRCNICI 309

SEQ ID NO: 108: H3-mini2-cl9+10+11

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHA VPNGT IVKTITNDQI 50
 EVTNATELVQ **SGGGG**RYVKQ NTLKLATGMR NVPEKQT**QGI** FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRL**R**GKTNEK 150
SHQTEKESSE GEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300
 GNIRCNICI 309

SEQ ID NO: 109: H3-mini2-cl9+10+11-tri

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHA VPNGT IVKTITNDQI 50
 EVTNATELVQ **SGGGG**RYVKQ NTLKLATGMR NVPEKQT**QGI** FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRL**R**GKTNEK 150
SHQTEKESSE GEGIEAIEKK IEAIEKKIEA IEKKELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300
 GNIRCNICI 309

SEQ ID NO: 110: H3-mini2-cl9+10+11-GCN4

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHA VPNGT IVKTITNDQI 50
 EVTNATELVQ **SGGGG**RYVKQ NTLKLATGMR NVPEKQT**QGI** FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRL**R**GKTNEK 150
SHQTEKESSE GEGRMKOIED KIEEIESKOK KIENELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300

GNIRCNICI 309

SEQ ID NO: 111: H3-mini2-cl9+10+11+12

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SGGGGRYVCQ NTLKLATGMR NVPEKQTQGI FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRLRGKTNEK 150
SHOTEKESSE GEGRIQDLEK YVEDTKIDLW CYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300

GNIRCNICI 309

SEQ ID NO: 112: H3-mini2-cl9+10+12

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SGGGGRYVCQ NTLKLATGMR NVPEKQTQGI FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRLLIGKTNEK 150
SHOTEKESSE GEGRIQDLEK YVEDTKIDLW CYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300

GNIRCNICI 309

SEQ ID NO: 113: H3-mini2-cl9+10+11+12-GCN4

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SGGGGRYVCQ NTLKLATGMR NVPEKQTQGI FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRLRGKTNEK 150
SHOTEKESSE GEGRMKOIED KIEEIESKLW CYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300

GNIRCNICI 309

SEQ ID NO: 114: H3-mini2-cl9+10+11+12-tri

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SGGGGRYVCQ NTLKLATGMR NVPEKQTQGI FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQINGKL NRLRGKTNEK 150
SHOTEKESSE GEGIEAIEKK IEAIEKKILW CYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300

GNIRCNICI 309

SEQ ID NO: 115: H3-mini2-cl9+13

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SGGGGRYVKQ NTLKLACGMR NVPEKQTQGI FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQCNGKL NRLLIGKTNEK 150
SHOTEKESSE VEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300

GNIRCNICI 309

SEQ ID NO: 116: H3-mini2-cl9+10+11+13

MKTIIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SGGGGRYVKQ NTLKLACGMR NVPEKQTQGI FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQCNGKL NRLRGKTNEK 150
SHOTEKESSE GEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300

GNIRCNICI 309

SEQ ID NO: 117: H3-mini2-cl9+10+11+13-GCN4

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SGGGGRYVKQ NTLKLACGMR NVPEKQTQGI FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQCNGKL NRLRGKTNEK 150
SHQTEKESSE GEGRMKOIED KIEEIESKOK KIENELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300
 GNIRCNICI 309

SEQ ID NO: 118: H3-mini2-cl9+10+11+13-tri

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SGGGGRYVKQ NTLKLACGMR NVPEKQTQGI FGAIAGFIEN 100
 GWEGMVDGWY GFRHQNSEGI GQAADLKSTQ AAINQCNGKL NRLRGKTNEK 150
SHQTEKESSE GEGIEAIEKK IEAIEKKIEA IEKKELLVAL ENQHTIDLTD 200
 SEMNKL FERT KKQLRENAED MGNGCFKIYH KCDNACIGSI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQK 300
 GNIRCNICI 309

SEQ ID NO: 119: H3-mini3-cl9+10+11+12+14

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSGGGGNDKP FQNVNRITYG AGPRYVCQNT LKLATGMRNV 100
 PEKQTQGIFG ALAGFIENGW EGMVDGWYGF RHQNSEGIGQ AADLKSTQAA 150
 INQINGKLNRLRGKTNEKSH QTEKESSEGE GRIQDLEKYV EDTKIDLWCY 200
 NAELLVALEN QHTIDLTDSE MNKL FERTKK QLRENAEDMG NGCFKIYHKC 250
 DNACIGSIRN GTYDHDVYRD EALNNRFQIK GVELKSGYKD WILWISFAIS 300
 CFLLCVLLG FIMWACQKGN IRCNICI 327

SEQ ID NO: 120: H3-mini4-cl9+10+11+12+14

MKTHIALSYI LCLVFAQKLP GNDNSTATLC LGHHAVPNGT IVKTITNDQI 50
 EVTNATELVQ SSSTGGGGYG AGPRYVCQNT LKLATGMRNV PEKQTQGIFG 100
 AIAGFIENGW EGMVDGWYGF RHQNSEGIGQ AADLKSTQAA INQINGKLN 150
 LRGKTNEKSH QTEKESSEGE GRIQDLEKYV EDTKIDLWCY NAELLVALEN 200
 QHTIDLTDSE MNKLFERTKK QLRENAEDMG NGCFKIYHKC DNACIGSIRN 250
 GTYDHDVYRD EALNNRFQIK GVELKSGYKD WILWISFAIS CFLLCVLLG 300
 FIMWACQKGN IRCNICI 317

SEQ ID NO: 121: H3 chiều dài đủ A/Hong Kong/1/1968

MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SSSTGKICNN PHRILDGIDC TLIDALLGDP HCDVFQNETW 100
 DLFVERSKAF SNCYPYDVPD YASLRSLVAS SGTLEFITEG FTWTGVTQNG 150
 GSNACKRGP GSGFFSRLNWL TKSGSTYPVL NVTMPNNDNF DKLYIWGVHH 200
 PSTNQEQTSL YVQASGRVTV STRRSQQTII PNIGSRPWVR GLSSRSIYW 250
 TIVKPGDVLV INSNGNLIAP RGYFKMRTGK SSIMRSDAPI DTCISECITP 300
 NGSIPNDKPF QNVNKITYGA CPKYVKQNTL KLATGMRNVP EKQTRGLFGA 350
 IAGFIENGWE GMIDGWYGF RHQNSEGTQA ADLKSTQAAI DQINGKLN 400
 IEKTNEKFHQ IEKEFSEVEG RIQDLEKYVE DTKIDLWSYN AELLVALENQ 450
 HTIDLTDSEM NKLFEKTRRQ LRENAEDMG NGCFKIYHKCD NACIESIRNG 500
 TYDHDVYRDE ALNNRFQIKG VELKSGYKDW ILWISFAISC FLLCVLLGF 550
 IMWACQRGNI RCNICI 566

SEQ ID NO: 122: HK68 H3m2-cl9

MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVKQ NTLKLATGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQINGKL NRVEKTNEK 150
 SHQTEKESSE VEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 123: HK68 H3m2-cl9+10

MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVKQ NTLKLATGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQINGKL NRVEKTNEK 150
 SHQTEKESSE GEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 124: HK68 H3m2-cl9+10+11

MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVKQ NTLKLATGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQINGKL NRVREKTNEK 150
 SHQTEKESSE GEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 125: HK68 H3m2-cl9+10+12
 MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVCQ NTLKLATGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQINGKL NRVIKTNEK 150
SHQTEKESSE GEGRIQDLEK YVEDTKIDLW CYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 126: HK68 H3m2-cl9+10+11+12
 MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVCQ NTLKLATGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQINGKL NRVREKTNEK 150
SHQTEKESSE GEGRIQDLEK YVEDTKIDLW CYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 127: HK68 H3m2-cl9+10+11+13
 MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVKQ NTLKLACGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQINGKL NRVREKTNEK 150
SHQTEKESSE GEGRIQDLEK YVEDTKIDLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 128: HK68 H3m2-cl9+10+11+12-tri
 MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVCQ NTLKLATGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQINGKL NRVREKTNEK 150
SHQTEKESSE GEGIEAIEKK IEAIEKKILW CYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 129: HK68 H3m2-cl9+10+11+13-tri
 MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVKQ NTLKLACGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQINGKL NRVREKTNEK 150
SHQTEKESSE GEGIEAIEKK IEAIEKKILW SYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 130: HK68 H3m2-cl9+10+11+12-GCN4

MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVCQ NTLKLATGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQINGKL NRVREKTNEK 150
SHQTEKESSE GEGRMKOIED KIEEIESKLW CYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 131: HK68 H3m2-cl9+10+11+13-GCN4

MKTHIALSYI FCLALGQDLP GNDNSTATLC LGHHAVPNGT LVKTITDDQI 50
 EVTNATELVQ SGGGGKYVKQ NTLKLACGMR NVPEKQTQGL FGAIAGFIEN 100
 GWEGMIDGWY GFRHQNSEGT GQAADLKSTQ AAIDQCNGKL NRVREKTNEK 150
SHQTEKESSE GEGRMKOIED KIEEIESKLW SYNAELLVAL ENQHTIDLTD 200
 SEMNKLFEKT RRQLRENAED MGNGCFKIYH KCDNACIESI RNGTYDHDVY 250
 RDEALNNRFQ IKGVELKSGY KDWILWISFA ISCFLLCVVL LGFIMWACQR 300
 GNIRCNICI 309

SEQ ID NO: 132: B/Florida/4/2006 HA chiều dài đủ

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50
 PTKSYFANLK GTRTRGKLCP DCLNCTDLDV ALGRPMC VGT TPSAKASILH 100
 EVKPVTS GCF PIMHDRTKIR QLPNLLRGYE NIRLSTQ NVI DAEKAPGGPY 150
 RLGTSGSCP N ATSKSGFFAT MAWAVPKDNN KNATNPLTVE VPYICTEGED 200
 QITVWGFHSD DKTQMKNLYG DSNPQKFTSS ANGVTTTHYVS QIGSFDPDQTE 250
 DGGLPQSGRI VVDYMMQKPG KTGTIVYQRG VLLPQK VWCA SGRSKVIKGS 300
 LPLIGEADCL HEKYGGLNKS KPYYTGEHAK AIGNCPIWVK TPLKLANGTK 350
 YRPPAKLLKE RGFFGAIAGF LEGGWEGMIA GWHGYTSHGA HGVAVAADLK 400
 STQEAINKIT KNLNSLSELE VKNLQRLSGA MDELHNEILE LDEKVDLLRA 450
 DTISSQIELA VLLSNEGIIN SEDEHLLALE RKLKMLGPS AVEIGNGCFE 500
 TKHKCNQTCL DRIAAGTFNA GEFSLPTFDS LNITAASLND DGLDNHTILL 550
 YYSTAASSLA VTLMLAIFIV YMVS RDNVSC SICL 584

SEQ ID NO: 133: FL4-06 B-m2

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50
GGGGIWVKTP LKLANGTKYR PPAKLLKEQG FFGAIAGFLE GGWEGM IAGW 100
 HGYTSHGAHG VAVAADLKST QEAINKITKN LNSLSELEVK NLQRLSGAMD 150
 ELHNEILELD EKVDDL RADT ISSQIELAVL LSNEGIINSE DEHLLALERK 200
 LKKMLGPSAV EIGNGCFETK HKCNQTCLDR IAGTFNAGE FSLPTFDSL N 250
 ITAASLNDDG LDNHTILLYY STAASSLAVT LMLAIFIVYM VSRDNVSCSI 300
 CL 302

SEQ ID NO: 134: FL4-06 B-m2-CL1+5

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50
GGGGIWVCTP LKLANGTKYR PPAKLLKEQG FFGAIAGFLE GGWEGM IAGW 100
 HGYTSHGAHG VAVAADLKST QEAINKITKN LNSLSELETK NSQRTSGAMD 150
EGHNEILELD EKVDDL RADT ICSQIELAVL LSNEGIINSE DEHLLALERK 200
 LKKMLGPSAV EIGNGCFETK HKCNQTCLDR IAGTFNAGE FSLPTFDSL N 250
 ITAASLNDDG LDNHTILLYY STAASSLAVT LMLAIFIVYM VSRDNVSCSI 300
 CL 302

SEQ ID NO: 135: FL4-06 B-m2-CL1+5-GCN4a

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50
GGGGIWVCTP LKLANGTKYR PPAKLLKEQG FFGAIAGFLE GGWEGM IAGW 100
 HGYTSHGAHG VAVAADLKST QEAINKITKN LNSLSELETK NSQRTSGAMD 150
EGHRRM KOIE DKIEEILSKI ICSQIELAVL LSNEGIINSE DEHLLALERK 200

LKKMLGPSAV EIGNGCFETK HKCNQTCLDR IAAGTFNAGE FSLPTFDSL 250
 ITAASLNDDG LDNHTILLY STAASSLAVT LMLAIFIVYM VSRDNVSCSI 300
 CL 302

SEQ ID NO: 136: FL4-06 B-m2-CL1+5-GCN4b

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50
GGGGIWVCTP LKLANGTKYR PPAKLLKEQG FFGALAGFLE GGWEGMIAGW 100
 HGYTSHGAHG VAVAADLKST QEAINKITKN LNSLSELETK NSQRTSGAMD 150
 EGHMKOIED KIEELSKIT ICSQIELAVL LSNEGIINSE DEHLLALERK 200
 LKKMLGPSAV EIGNGCFETK HKCNQTCLDR IAAGTFNAGE FSLPTFDSL 250
 ITAASLNDDG LDNHTILLY STAASSLAVT LMLAIFIVYM VSRDNVSCSI 300
 CL 302

SEQ ID NO: 137: B/Malaysia/2506/2004 HA chiều dài đủ

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50
 PTKSHFANLK GTETRGLCP KCLNCTDLV ALGRPKCTGN IPSARVSILH 100
 EVRPVTS GCF PIMHDRTKIR QLPNLLRGYE HIRLSTHNVI NAENAPGGPY 150
 KIGTSGSCP VTNNGGFFAT MAWAVPKNDN NKTATNSLTI EYPYICTEGE 200
 DQITVWGFHS DNEAQMAKLY GDSKPQKFTS SANGVTTHYV SQIGGFNPQT 250
 EDGGLPQSGR IVVDYMVQKS GKTGTITYQR GILLPQKVWC ASGRSKVIK 300
 SLPLIGEADC LHEKYGGLNK SKPYTGEHA KAIGNCPIWV KTPLKLANGT 350
 KYRPPAKLLK ERGFFGALAG FLEGGWEGMI AGWHGYTSHG AHGVAVAADL 400
 KSTQEAINKI TKNLNSLSEL EVKNLQRLSG AMDELHNEIL ELDEKVDDL 450
 ADTISSQIEL AVLLSNEGII NSEDEHLLAL ERKLKMLGP SAVEIGNGCF 500
 ETKHKCNQTC LDRIAAGTFD AGEFSLPTFD SLNITAASLN DDGLDNHTIL 550
 LYSTAASSL AVTLMIAIFV VYMVSRDNVS CSICL 585

SEQ ID NO: 138: Mal2506-04 B-m2

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50
GGGGIWVKTP LKLANGTKYR PPAKLLKEQG FFGALAGFLE GGWEGMIAGW 100
 HGYTSHGAHG VAVAADLKST QEAINKITKN LNSLSELEVK NLQRLSGAMD 150
 ELHNEILELD EKVDDLADT ISSQIELAVL LSNEGIINSE DEHLLALERK 200
 LKKMLGPSAV EIGNGCFETK HKCNQTCLDR IAAGTFDAGE FSLPTFDSL 250
 ITAASLNDDG LDNHTILLY STAASSLAVT LMIAIFVYVYM VSRDNVSCSI 300
 CL 302

SEQ ID NO: 139: Mal2506-04 B-m2-CL1+5

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50
GGGGIWVCTP LKLANGTKYR PPAKLLKEQG FFGALAGFLE GGWEGMIAGW 100
 HGYTSHGAHG VAVAADLKST QEAINKITKN LNSLSELETK NSQRTSGAMD 150
 EGHNEILELD EKVDDLADT ICSQIELAVL LSNEGIINSE DEHLLALERK 200
 LKKMLGPSAV EIGNGCFETK HKCNQTCLDR IAAGTFDAGE FSLPTFDSL 250
 ITAASLNDDG LDNHTILLY STAASSLAVT LMIAIFVYVYM VSRDNVSCSI 300
 CL 302

SEQ ID NO: 140: Mal2506-04 B-m2-CL1+5-GCN4a

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50
GGGGIWVCTP LKLANGTKYR PPAKLLKEQG FFGALAGFLE GGWEGMIAGW 100
 HGYTSHGAHG VAVAADLKST QEAINKITKN LNSLSELETK NSQRTSGAMD 150
 EGHRRMKOIE DKIEELSKI ICSQIELAVL LSNEGIINSE DEHLLALERK 200
 LKKMLGPSAV EIGNGCFETK HKCNQTCLDR IAAGTFDAGE FSLPTFDSL 250
 ITAASLNDDG LDNHTILLY STAASSLAVT LMIAIFVYVYM VSRDNVSCSI 300
 CL 302

SEQ ID NO: 141: Mal2506-04 B-m2-CL1+5-GCN4b

MKAIIVLLMV VTSNADRICT GITSSNSPHV VKTATQGEVN VTGVIPLTTT 50

GGGGIWVCTP LKLANGTKYR PPAKLLKE**QG** FFGA**IAGFIE** GGWEGM**IAGW** 100
 HGYTSHGAHG VAVAADLKST QEAINKITKN LNSLSELE**TK** NS**QRT**SGAMD 150
EGHRMKOIED **KIEEILSKIT** ICSQIELAVL LSNEGIINSE DEHLLALERK 200
 LKKMLGPSAV EIGNGCFETK HKCNQTCLDR IAAGTFDAGE FSLPTFDSL 250
 ITAASLNDDG LDNHTILLYY STAASSLAVT LMIAIFVVYM VSRDNVSCSI 300
 CL 302

SEQ ID NO: 142: Trình tự gốc liên ứng của HA của virus cúm B 416-436

416 LSELEVKNLQRLSGAMDELHN **436**

SEQ ID NO: 144: s-H1-mini2-cụm1+5+6-trim (A/Brisbane/59/2007)
 MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENG**GGGGK**YVC SAKLRMTGL RNIPSIQ**SQG** LFGA**IAGFIE** GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT Q**STATGKEGN** 150
KSEIEAIEKK IEAIEKKIEI WCYN**AELLVL** LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVQIEGRHH HHHHH 275

SEQ ID NO: 145: s-H1-mini2-cụm1+5+6-GCN4 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENG**GGGGK**YVC SAKLRMTGL RNIPSIQ**SQG** LFGA**IAGFIE** GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT Q**STATGKEGN** 150
KSERMKOIED **KIEEIESKOI** WCYN**AELLVL** LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVQIEGRHH HHHHH 275

SEQ ID NO: 146: s-H1-mini2-cụm1+5+6 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENG**GGGGK**YVC SAKLRMTGL RNIPSIQ**SQG** LFGA**IAGFIE** GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT Q**STATGKEGN** 150
KSERRMENLN KKVDDGFIDI WCYN**AELLVL** LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVQIEGRHH HHHHH 275

SEQ ID NO: 147: s-H1-mini2-cụm1+5+6 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENG**GGGGK**YVC SAKLRMTGL RNIPSIQ**SQG** LFGA**IAGFIE** GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT Q**STATGKEGN** 150
KSERRIENLN KKIDDGFI**DI** WCYN**AELLVL** LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVQIEGRHH HHHHH 275

SEQ ID NO: 148: s-H1-mini2-cum1+5 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENGGGGKYVC SAKLRMVTGL RNIPSIQSQG LFGAIAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEFN 150
 KSERRMENLN KKVDDGFIDI WCYNAELLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVYQIEGRHH HHHHH 275

SEQ ID NO 149: s-H1 Chiều dài đủ R343Q (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENSHNGKLCL LKGIAPLQLG NCSVAGWILG NPECELLISK ESWSYIVEKP 100
 NPENGTCTPG HFADYEELRE QLSSVSSFER FEIFPKESSW PNHTVTGVSA 150
 SCSHNGESSF YRNLLWLTGK NGLYPNLSKS YANNKEKEVL VLWGVHHPN 200
 IGDQKALYHT ENAYVSVVSS HYSRKFTPEI AKRPKVRDQE GRINYYWTLL 250
 EPGDTIIFEA NGNLIAPRYA FALSRRGFGSG IINSNAPMDK CDAKCQTPQG 300
 AINSSLPFQN VHPVTIGCEP KYVRSALRM VTGLRNIPSI QSQGLFGAIA 350
 GFIEGGWTGM VDGWYGYHHQ NEQSGYAAD QKSTQNAING ITNKVNSVIE 400
 KMNTQFTAVG KEFNKLERRM ENLNKKVDDG FIDIWTYNAE LLVLENERT 450
 LDFHDSNVKN LYEKVKSQLK NNAKEIGNGC FEFYHKCNDE CMESVKNGTY 500
 DYPKYSEESK LNREKIDGVK LESMGVVYQIE GRHHHHHHH 539

SEQ ID NO: 150: s-H1-mini2-cum1+5+6-nl (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENHNGKKYVC SAKLRMVTGL RNIPSIQSQG LFGAIAGFIE GGWTGMVDGW 100
 YGYHHQNEQG SGYAADQKST QNAINGITNK VNSVIEKMNT QSTATGKEGN 150
 KSERRMENLN KKVDDGFIDI WCYNAELLVL LENERTLDFH DSNVKNLYEK 200
 VKSQLKNNAK EIGNGCFEFY HKCNDECMES VKNGTYDYPK YSEESKLNRE 250
 KIDGVKLESM GVVYQIEGRHH HHHHH 275

SEQ ID NO: 151: s-H1-mini2-cum1+5+6-nl2 (A/Brisbane/59/2007)

MKVKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 50
 ENHNGKYVCS AKLRMVTGLR NIPSIQSQGL FGALAGFIEG GWTGMVDGWY 100
 GYHHQNEQGS GYAADQKSTQ NAINGITNKV NSVIEKMNTQ STATGKEGNK 150
SERRMENLNK KVDDGFIDIW CYNAELLVLL ENERTLDFHD SNVKNLYEKV 200
 KSQLKNNAKE IGNCGCFEFYH KCNDECMESV KNGTYDYPKY SEESKLNREK 250
 IDGVKLESMG VYQIEGRHHH HHHH 274

SEQ ID NO:152 H1mini2a-cl1+5+6_ không cầu nối(HNGK)

MKVKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLL 50
 ENHNGKKYVCSAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGW 100
 YGYHHQNEQSGSYAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGN 150
 KSERRMENLNKKVDDGFIDIWCYNAELLVLENERLDFHDSNVKNLYEK 200
 VKSQLKNNAKEIGNGCFEFYHKCNDECMESVKNGTYDYPKYSEESKLNRE 250
 KIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC 300

I 301

SEQ ID NO:153 H1mini2a-cl1+5+6_không_cầu_nối2s

MKVKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTTHSVNLL 50
 ENHNGKYVCSAKLRMVTGLRNIPSIQSQGLFGAIAAGFIEGGWTGMVDGWY 100
 GYHHQNEQGSYGAAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGNK 150
 SERRMENLNKKVDDGFIDIWCYNAELLVLENERTLDFHDSNVKNLYEKV 200
 KSQLKNNAKEIGNGCFEFYHKCNDECMESVKNGTYDYPKYSEESKLNREK 250
 IDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC 300

SEQ ID NO:154 H1-mini2-cl1+5+6-no_cầu_nối2s-GCN4

MKVKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTTHSVNLL 50
 ENSHNGKYVCSAKLRMVTGLRNIPSIQSQGLFGAIAAGFIEGGWTGMVDGW 100
 YGYHHQNEQGSYGAAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGN 150
 KSERMKQIEDKIEEIESKQIWCYNAELLVLENERTLDFHDSNVKNLYEK 200
 VKSQLKNNAKEIGNGCFEFYHKCNDECMESVKNGTYDYPKYSEESKLNRE 250
 KIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC 300

I 301

SEQ ID NO:155 H1mini2a-cl1+5+6_không_cầu_nối2s-trim3

MKVKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTTHSVNLL 50
 ENSHNGKYVCSAKLRMVTGLRNIPSIQSQGLFGAIAAGFIEGGWTGMVDGW 100
 YGYHHQNEQGSYGAAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGN 150
 KSERRIEAIEKKIEAIEKKIWCYNAELLVLENERTLDFHDSNVKNLYEK 200
 VKSQLKNNAKEIGNGCFEFYHKCNDECMESVKNGTYDYPKYSEESKLNRE 250
 KIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC 300

I 301

SEQ ID NO:156 H1mini2a-cl1+5+6-12

MKVKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTTHSVNLL 50
 ENGGGGKYVCSAKLRMVTGLRNNPSNQSQGLFGAIAAGYIEGGWTGMVDGW 100
 YGYHHQNEQGSYGAAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGN 150
 KSERRMENLNKKVDDGFIDIWCYNAELLVLENERTLDFHDSNVKNLYEK 200
 VKSQLKNNAKEIGNGCFEFYHKCNDECMESVKNGTYDYPKYSEESKLNRE 250
 KIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC 300

I 301

SEQ ID NO:157 H1mini2a-cl1+5+6-12+13

MKVKLLVLLCTFTATYADTICIGYHANNSTDTVDTVLEKNVTVTTHSVNLL 50
 ENGGGGKYVCSAKLRMVTGLRNNPSNQSQGLFGAIAAGYIEGGWTGMVDGW 100
 YGYHHQNEQGSYGAAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGN 150
 KSERRMENLNKKVDDGFIDIWCYNAELLVLENERTLDFHDSNVKNLYEK 200
 VKSQLKNNAKEIGNGCFEFYHKCNDECMESVKNGTYDYPKYSEESKLNRE 250
 KIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC 300

SEQ ID NO:158:H5 FL HA A/Vietnam/1203/2004

(341 RRRKK 345 được loại bỏ và đột biến R346Q được đưa vào)

mekivllfaivslvksdqicigyhannsteqvdtmeknvtvthaqdile 50
 kkhngklcdldgvkpllrdesvagwllgnpmcdefinwpewsyivekan 100
 pvndlcypgdfndyeelkhllsrinhfekiqiipksswssheaslgvssa 150
 cpyqgkssffrnvwlikknstyptikrsynntnqedllvlgihhpnda 200
 aeqtlyqnpttysvgtstlnqrlvpriatrskvngqgrmeffwtilk 250
 pndainfesngnfiapeyaykivkkgdstimkseleygncntkcqtpmga 300
 inssmpfhnhptligecpkvksnrlvlatglrnspqrerrrrkrglfg 350
 aiagfieggwqgmvdgwygyhhsneqsgyaadkestqkaidgvtnkvns 400
 iidkmntqfeavgrefnnlerrienlnkkmedgflvwtynaellvmlen 450
 ertldfhdsnvknlyekvrsqknnakeigngcfehyhkcdnecmesvrn 500
 gtydypkyseeaklnreeidgvklestriyqilaiystvasslvvlvslgai 550
 glslwmcsngslqcrici 568

SEQ ID NO:159:H1 FL HA A/California/04/2009 R343Q

MKAILVLLYTFATANADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLL 50
 EDKHNGKLCKLRGVAPLHLGKCNIAAGWILGNPECESLSTASSWSYIVETP 100
 SSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVT 150
 AACPHAGAKSFYKNLIWL VKKGNSYPKLSKSYINDKGKEVLVLWGIHPS 200
 TSADQQSLYQNADTYVFGSSRYSKKFKPEIAIRPKVRDQEGRMNYYWTL 250
 VEPGDKITFEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPK 300
 GAINSLPQNIHPITIGKCPKYVKSTKLRLATGLRNIPSIQSRGLFGAI 350
 AGFIEGGWTGMVDGWYGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVI 400
 EKMNTQFTAVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLENER 450
 TLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGT 500
 YDYPKYSEEAKLNREEIDGVKLESTRIYQILAIYSTVASSLVVLVSLGAI 550
 SFWMCSNGSLQCRICI 566

SEQ ID NO:160:H1mini-HAA/California/07/2009

MKAILVLLYTFATANADTLCIGYHANNSTDTVDTVLEKNVTVTHSVNLL 50
 EDGGGGKYVCSTKLRLATGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGW 100
 YGYHHQNEQSGYAADLKSTQNAIDEITNKVNSVIEKMNTQSTATGKEGN 150
 HSERMKQIEDKIEEIESKQIWCYNAELLVLENER TLDYHDSNVKNLYEK 200
 VRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGTYDYPKYSEEAKLNRE 250
 EIDGVKLESTRIYQILAIYSTVASSLVVLVSLGAISFWMCSNGSLQCRIC 300

I30 1

SEQ ID NO:161:H1mini-HAA/PuertoRico/8/1934

MKANLLVLLCALAAADADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLL 50
 EDGGGGKYVCSAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMIDGW 100
 YGYHHQNEQSGYAADQKSTQNAINGITNKVNTVIEKMNTQSTATGKEGN 150
 KSERMKQIEDKIEEIESKQIWCYNAELLVLENER TLDYHDSNVKNLYEK 200
 VKSQLKNNAKEIGNGCFEFYHKCDNECMESVRNGTYDYPKYSEESKLNRE 250
 KVDGVKLESMGIYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC 300

I301

SEQ ID NO:162:H1mini-HAA/Texas/36/1991

MKAKLLVLLCAFTATYADTICIGYHANNSTDTVDTVLEKNVTVTTHSVNLL 50
 EDGGGGKYVCSTKLRMVTGLRNIPSIQSQGLFGAIAAGFIEGGWTGMIDGW 100
 YGYHHQNEQSGYAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGN 150
 KSERMKQIEDKIEEIESKQIWCYNAELLVLLENGRTLDFHDSNVKNLYEK 200
 VKSQLKNNAKEIGNGCFEFYHKCNNECMESVKNGTYDYPKYSEESKLNRG 250
 KIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC 300
 I301

SEQ ID NO:163:H5mini-HAA/Vietnam/1203/2004

MEKIVLLFAIVSLVKSQICIGYHANNSTEQVDTIMEKNVTVTTHAQDILE 50
 KGGGGKYVCNRLVLATGLRNSPQRESQGLFGAIAAGFIEGGWQGMVDGWY 100
 GYHHSNEQSGYAADKESTQKAIDGVTNKVNSIIDKMNTQSEATGREGNN 150
 SERMKQIEDKIEEIESKQIWCYNAELLVLMENERTLDFHDSNVKNLYDKV 200
 RLQLRDNAKELGNGCFEFYHKCDNECMESVRNGTYDYPQYSEEARLKREE 250
 ISGVKLESIGIYQILSIYSTVASSLALAIMVAGLSLWMCSNGSLQCRICI 300

SEQ ID NO:164: mHA_H1N1_A_Maryland_12_1991

MKAILLVLLYTFTAANADTLCIGYHANNSTDTVDTVLEKNVTVTTHSVNLLEDGGG
 KYVC
 STKLRLMATGLRNIPSIQSQGLFGAIAAGFIEGGWTGMIDGWYGYHHQNEQSGYAAD
 QKST
 QNAIDGITNKVNSVIEKMNTQSTATGKEGNHSERMKQIEDKIEEIESKQVWCYNAEL
 LVL
 LENERTLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDDTCMESVKNGT
 YDYPK
 YSEESKLNREEIDGVKLESTRIYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRI
 C
 I*

SEQ ID NO:165: mHA_H1N1_A_Henry_1936

MKARLLVLLCALAATDADTICIGYHANNSTDTVDTVLEKNVTVTTHSVNLLEDGGG
 GKYYC
 SAKLRMTGLRNIPSIQSQGLFGAIAAGFIEGGWTGMIDGWYGYHHQNEQSGYAAD
 QKST
 QNAINGITNKVNSVIEKMNTQSTATGKEGNNSERMKQIEDKIEEIESKQIWCYNAEL
 LVL
 LENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCDNECMESVRNGT
 YDYPK
 YSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQC
 RIC
 I*

SEQ ID NO:166: mHA_H1N1_A/AA/Marton/1943

MKARLLVLLCALAATDADTICIGYHANNSTDTVDTVLEKNVTVTTHSVNLLEDGGG
 GKYYC
 STKLRLMTGLRNIPSIQSQGLFGAIAAGFIEGGWTGMIDGWYGYHHQNEQSGYAAD
 QKST
 QNAINGITNKVNSVIEKMNTQSTATGKEGNNSERMKQIEDKIEEIESKDIWCYNAEL
 LVL
 LENERTLDFHDSNVKNLYEKVKNQLRNNAKEIGNGCFEFYHKCNNECMESVKNGT
 YDYPK
 YSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQC
 RIC

I

SEQ_ID_NO: 167: mHA_H1N1_A New_York_607_1995

MKAKLLVLLCAFTATYADTICIGYHANNSTDVDTVLEKNVTYTHSVNLLEDGGGG
KYVCSTKLRMVTGLRNIPSIQSQGLFGAIAAGFIEGGWTGMIDGWYGYHHQNEQSGYAAD
QKSTQNAIDGITNKVNSVIEKMNTQSTATGKEGNKSERMKQIEDKIEEIESKQIWCYNAEL
LVLLENERTLDFHDSNVKNLYEKVKTQLKNNAKEIGNGCFEFYHKCNNECMESVKNGT
YDYPKYSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQC
RIC

I*

SEQ_ID_NO: _168: mHA_H1N1_A_New_Jersey_11_2007

1

MKARLLVLLCALAATDADTICIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDGGGGKY
VC

61

SAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADQ
KST

121

QNAINGITNKVNSVIEKMNTQSTATGKEGNKSERMKQIEDKIEEIESKQIWCYNAELLVL
181

LENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCNDECMESVKNGTYD
YPK

241

YSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC
301 I*

SEQ_ID_NO: _169: mHA_H1N1_A_USSR_92_1977

1

MKAKLLVLLCALSATDADTICIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDGGGGKY
VC

61

STKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQSGYAADQK
ST

121

QNAINGITNKVNSVIEKMNTQSTATGKEGNKSERMKQIEDKIEEIESKQIWCYNAELLVL
181

LENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCNNECMESVKNGTYD
YPK

241

YSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC
301 I*

SEQ_ID_NO: _170: mHA_H1N1_A_New_York_629_1995

1

MKVKLLVLLCAFTATYADTICIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDGGGGKY
VC

61

STKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQSGYAADQK
ST

121

QNAIDGITNKVNSVIEKMNTQSTATGKEGNKSERMKQIEDKIEEIESKQIWCYNAELLVL
181

LENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCNNECMESVKNGTYD
YPK

241

YSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCRIC
301 I*

SEQ_ID_NO: _171: mHA_H1N1_A_Virginia_UR06-0549_2007

1MKVKLLVLLCTFTATYADTICIGYHANNSTDVDTVLEKNVTVTHSVNLLLEDGGGGK
YVC

61SAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAAD
QKST

121QNAINGITNKVNSVIEKMNTQSTATGKEGNKSERMKQIEDKIEEIESKQIWCYNAELL
VL

181LENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCNDECMESVKNGT
YDYPK

241YSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVSLGAISFWMCSNGSLQCR
IC

301I*

SEQ_ID_NO: _172: _mHA_H1N1_A_Texas_UR0-0526_2007

1

MKVKLLVLLCTFTATYADTICIGYHANNSTDVDTVLEKNVTVTHSVNLLEDGGGGKY
VC

61

SAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADQ
KST

121

QNAINGITNKVNSVIEKMNTQSTATGKEGNKSERMKQIEDKIEEIESKQIWCYNAELLVL
181

LENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCNDECMESVKNGT
YDYPK

241

YSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLISLGAISFWMCSNGSLQCRIC

301 I*

SEQ_ID_NO: _173: _mHA_H1N1_A_Sydney_DD3-55_2010

1

MKAILVLLYTFATANADTLCIGYHANNSTDVDTVLEKNVTVTHSVNLLEDGGGGKY
VC

61

STKLRLATGLRNVPSIQSQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQSGYAADLK
ST

121

QNAIDEITNKVNSVIEKMNTQSTATGKEGNHSERMKQIEDKIEEIESKQIWCYNAELLVL
181

LENERTLDYHDSNVKNLYEKVRSQKNNAKEIGNGCFEFYHKCDNTCMESVKNGT
YDYPK

241

YSEEAKLNREEIDGVKLESTRIYQILAIYSTVASSLVLVVSLGAISFWMCSNGSLQCRIC

301 I*

SEQ_ID_NO: 174 H3mini2a-cầu nối+cl9_+10+11+12+GCN4T_CG7-1
(A/HongKong/1/1968(H3N2))

1
MKTIIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNGTLVKTTITDDQIEVTNATELVQ
61
SGGGGKYVCQNTLKLATGMRNVPEKQTQGLFGAIAGFIENGWEGMIDGWYGFRHQNS
EGT
121
GQAADLKSTQAAIDQINGKLN RVREKTNEKSHQTEKESSNATGRMKQIEDKIEEIESKL
W
181
CYN AELLVALENQHTIDLT DSEM NKLFEKTRRQLRENAEDMGNGCFKIYHKCDNACIES
I
241
RNGTYDHDVYRDEALNNRFQIKGV ELKSGYKDWILWISFAISCFLLCVVLLGFIMWACQ
R
301 GNIRCNICI

SEQ_ID_NO: 175 H3mini2a-cầu nối+cl9_+10+12+18+GCN4T
(A/HongKong/1/1968(H3N2))

1
MKTIIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNGTLVKTTITDDQIEVTNATELVQ
61
SGGGGKYVCQNTLKLATGMRNVPEKQTQGLFGAIAGFIENGWEGMIDGWYGFRHQNS
EGT
121
GQAADLKSTQAAIDQINGKLN RVIEKTNEKSHQTEKESSEGE GNATGGCCGGRMKQIED
K
181
IEEIESKLWCYN AELLVALENQHTIDLT DSEM NKLFEKTRRQLRENAEDMGNGCFKIYH
K
241
CDNACIESIRNGTYDHDVYRDEALNNRFQIKGV ELKSGYKDWILWISFAISCFLLCVVLL
301 GFIMWACQ RGNIRCNICI

SEQ_ID_NO: 176 H3mini2a-cầu nối+cl9_+10+12+16+CG7-GCN4T
(A/HongKong/1/1968(H3N2))]

1
MKTIIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNGTLVKTTITDDQIEVCNATELVQ
61
SGGGGKYVCQNTLKLATCMRN VPEKQTQGLFGAIAGFIENGWEGMIDGWYGFRHQNS
EGT
121
GQAADLKSTQAAIDQINGKLN RVIEKTNEKSHQTEKESSNATGRMKQIEDKIEEIESKLW
181
CYN AELLVALENQHTIDLT DSEM NKLFEKTRRQLRENAEDMGNGCFKIYHKCDNACIES
I
241
RNGTYDHDVYRDEALNNRFQIKGV ELKSGYKDWILWISFAISCFLLCVVLLGFIMWACQ
R
301 GNIRCNICI
SEQ_ID_NO: 177 H3mini2a-cầu nối+cl9_+10+12+19+GCN4T

(A/HongKong/1/1968(H3N2))]

1

MKTIIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNGTLVKTTITDDQIEVTNATELVQ

61

SGGGGKYVCQNTLKLATGMRNVPEKQTQGLFGAAGFIENGWEGMIDGWYGFRHQNS

EGT

121

GQAADLKSTQAAIDQINGKLN RVIEKTNEKSHQTEKESSEGE GSGSGGCCGGRMKQIED

K

181

IEEIESKLWCYNAELLVALENQHTIDLT DSEMKNLF EKTRRQLRENAEDMGNGCFKIYH

K

241

CDNACIESIRNGTYDHDVYRDEALNNRFQIKGV ELKSGYKDWILWISFAISCFLLCVVLL

301 GFIMWACQQRGNIRCNICI

SEQ_ID_NO: _178 H3mini2a-cầu nối+cl9_+10+12+17+CG7-GCN4T

(A/HongKong/1/1968(H3N2))]

1

MKTIIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNCTLVKTTITDDQICVTNATELVQ

61

SGGGGKYVCQNTLKLATGMRNVPEKQTQGLFGAAGFIENGWEGMIDGWYGFRHQNS

EGT

121

GQAADLKSTQAAIDQINGKLN RVIEKTNEKSHQTEKESSNATGRMKQIEDKIEEIESKLW

181

CYNAELLVALENQHTIDLT DSEMKNLF EKTRRQLRENAEDMGNGCFKIYHKCDNACIES

I

241

RNGTYDHDVYRDEALNNRFQIKGV ELKSGYKDWILWISFAISCFLLCVVLLGFIMWACQ

R

301 GNIRCNICI

SEQ_ID_NO: _179 H3_HK68_mini2a-cầu nối2+cl9_+10+12+GCN4T

1

LATMKTIIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNGTLVKTTITDDQIEVTNATE

61

LVQSGSGSGGKYVCQNTLKLATGMRNVPEKQTQGLFGAAGFIENGWEGMIDGWYGF

RHQ

121

NSEGTGQAADLKSTQAAIDQINGKLN RVIEKTNEKSHQTEKESSEGEGRMKQIEDKIEEI

181

ESKLWCYNAELLVALENQHTIDLT DSEMKNLF EKTRRQLRENAEDMGNGCFKIYHKCD

NA

241

CIESIRNGTYDHDVYRDEALNNRFQIKGV ELKSGYKDWILWISFAISCFLLCVVLLGFIM

301 WACQQRGNIRCNICI**

SEQ_ID_NO: _180 H1-mini2-cụm1+5+6+GCN4-T49N

MKVKLLVLLCTFTATYADTICIGYHANNSTD TVD TVLEKNVT VTHSVNLL ENGGGG

KYVCSAKLRMVTGLRNIPSIQSQGLFGAAGFIEGGWTGMVDGWYGYHHQNEQGS

GYAADQKSTQNAINGINN KVN SVIEKMNTQSTATGKEGNKSERMKQIEDKIEEIESK

QIWCYNAELLV LLENERTLDFHDSNVKNLYEKVKSQ LKNNAKEIGNGCFEFYHKCN

DECMESVKNGTYDYPKYSEESKLNREKIDGVKLESMGVYQILAIYSTVASSLVLLVS
LGAISFWMCSNGLQCRICI

SEQ ID NO: 181 sH1-mini2-cl1+5+6-GCN4-Bromelain

MKVKLLVLLCTFTATYADTICIGYHANNSTDVDTVLEKNVTVTTHSVNLLENG
GGKGYVCSAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQG
SGYAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGNKSERMKQIEDKIEEIES
KQIWCYNAELLVLLNERTLDFHDSNVKNLYEKVKSQKNNAKEIGNGCFEFYHKC
NDECMESVKNGTYDYPKYSEESKLNREKIDGRSLVPRGSPGHHHHHH

SEQ ID NO: 182 sH1-mini2-cl1+5+6-GCN4-Bromelain-Foldon

MKVKLLVLLCTFTATYADTICIGYHANNSTDVDTVLEKNVTVTTHSVNLLENG
GGKGYVCSAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQG
SGYAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGNKSERMKQIEDKIEEIES
KQIWCYNAELLVLLNERTLDFHDSNVKNLYEKVKSQKNNAKEIGNGCFEFYHKC
NDECMESVKNGTYDYPKYSEESKLNREKIDGRSLVPRGSPGSGYIPEAPRDGQAYVR
KDGEWVLLSTFLGHHHHHH

SEQ ID NO: 183 sH1-mini2-cl1+5+6-GCN4t2

MKVKLLVLLCTFTATYADTICIGYHANNSTDVDTVLEKNVTVTTHSVNLLENG
GGKGYVCSAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQG
SGYAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGNKSERRMKQIEDKIEEIE
SKIWCYNAELLVLLNERTLDFHDSNVKNLYEKVKSQKNNAKEIGNGCFEFYHKC
NDECMESVKNGTYDYPKYSEESKLNREKIDGVKLESMGVYQIEGRHHHHHHHH

SEQ ID NO: 184 sH1-mini2-cl1+5+6-GCN4t2-Bromelain

MKVKLLVLLCTFTATYADTICIGYHANNSTDVDTVLEKNVTVTTHSVNLLENG
GGKGYVCSAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQG
SGYAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGNKSERRMKQIEDKIEEIE
SKIWCYNAELLVLLNERTLDFHDSNVKNLYEKVKSQKNNAKEIGNGCFEFYHKC
NDECMESVKNGTYDYPKYSEESKLNREKIDGRSLVPRGSPGHHHHHH

SEQ ID NO: 185 sH1-mini2-cl1+5+6-GCN4t2-Bromelain-Foldon

MKVKLLVLLCTFTATYADTICIGYHANNSTDVDTVLEKNVTVTTHSVNLLENG
GGKGYVCSAKLRMVTGLRNIPSIQSQGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQG
SGYAADQKSTQNAINGITNKVNSVIEKMNTQSTATGKEGNKSERRMKQIEDKIEEIE
SKIWCYNAELLVLLNERTLDFHDSNVKNLYEKVKSQKNNAKEIGNGCFEFYHKC
NDECMESVKNGTYDYPKYSEESKLNREKIDGRSLVPRGSPGSGYIPEAPRDGQAYVR
KDGEWVLLSTFLGHHHHHH

SEQ ID NO: 186 sH3 HK mini2a-cầu nối+cl9 +10+11+12+GCN4T-CG7-His

MKTHIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNGTLVKTITDDQIEVTNATE
LVQSGGGGKGYVCQNTLKLATGMRNVPEKQTQGLFGAIAGFIENGWEGMIDGWYGF
RHQNSEGTGQAADLKSTQAAIDQINGKLNVRREKTNEKSHQTEKESSNATGRMKQI
EDKIEEIESKLWCYNAELLVALENQHTIDLTDSEMKNLFKTRRQLRENAEDMGNG
CFKIYHKCDNACIESIRNGTYDHDVYRDEALNNRFQIKGRSLVPRGSPGHHHHHH

SEQ ID NO: 187 sH3 HK mini2a-cầu nối+cl9 +10+11+12+GCN4T-CG7-Foldon-His

MKTHIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNGTLVKTITDDQIEVTNATE
LVQSGGGGKGYVCQNTLKLATGMRNVPEKQTQGLFGAIAGFIENGWEGMIDGWYGF
RHQNSEGTGQAADLKSTQAAIDQINGKLNVRREKTNEKSHQTEKESSNATGRMKQI
EDKIEEIESKLWCYNAELLVALENQHTIDLTDSEMKNLFKTRRQLRENAEDMGNG
CFKIYHKCDNACIESIRNGTYDHDVYRDEALNNRFQIKGRSLVPRGSPGSGYIPEAPR
DGQAYVRKDGEWVLLSTFLGHHHHHH

YÊU CẦU BẢO HỘ

1. Polypeptit vùng gốc hemagglutinin của virus cúm chứa (a) đoạn đầu tận cùng đầu N của H1 HA1 chứa các axit amin 1-52 được liên kết cộng hóa trị bằng trình tự liên kết gồm 0-10 gốc axit amin với đoạn gốc đầu tận cùng đầu C của H1 HA1 chứa các axit amin 321-điểm cuối của HA1, và (b) vùng HA2 của hemagglutinin của virus cúm, trong đó gốc axit amin đầu tận cùng đầu C của đoạn gốc đầu tận cùng C của HA1 là glutamin (Q), và trong đó axit amin ở vị trí 406, 409, 413 và 416 được thay đổi thành axit amin được chọn từ nhóm bao gồm S, T, Q, và G, và trong đó các polypeptit chứa cầu disulphit giữa các axit amin ở vị trí 324 và 436; trong đó polypeptit chứa vùng gốc hemagglutinin từ HA của virus cúm A/Brisbane/59/2007 chứa trình tự nêu trong SEQ ID NO:1 hoặc vùng gốc hemagglutinin từ HA của virus cúm A/California/04/2009 chứa trình tự nêu trong SEQ ID NO:159; và

trong đó việc đánh số dựa trên việc đánh số của axit amin của chủng cúm H1N1 của trình tự nêu trong SEQ ID NO:1, và

trong đó polypeptit liên kết chọn lọc với kháng thể CR6261 và CR9114 chứa vùng CDR1 chuỗi nặng có trình tự nêu trong SEQ ID NO:25, vùng CDR2 chuỗi nặng có trình tự nêu trong SEQ ID NO:26, vùng CDR3 chuỗi nặng có trình tự SEQ ID NO:27, và vùng CDR1 chuỗi nhẹ có trình tự nêu trong SEQ ID NO: 28, vùng CDR2 chuỗi nhẹ có trình tự nêu trong SEQ ID NO:29 và vùng CDR3 chuỗi nhẹ có trình tự nêu trong SEQ ID NO:30, và không liên kết với kháng thể CR8057.

2. Polypeptit theo điểm nêu trên, trong đó polypeptit này không chứa vùng HA1 có chiều dài đủ.

3. Polypeptit theo điểm 1 hoặc 2, trong đó polypeptit được glycosyl hóa.

4. Polypeptit theo điểm 1, 2 hoặc 3, trong đó trình tự IEAIEKKIEAIEKKIE (SEQ ID NO: 83) được đưa vào ở vị trí 418 đến 433 hoặc trong đó trình tự MKQIEDKIEEIESKQ (SEQ ID NO: 84) được đưa vào ở vị trí 419 đến 433.

5. Polypeptit theo điểm 1, 2 hoặc 3, trong đó các axit amin M420, L423, V427, G430 được gây đột biến thành isoleuxin.

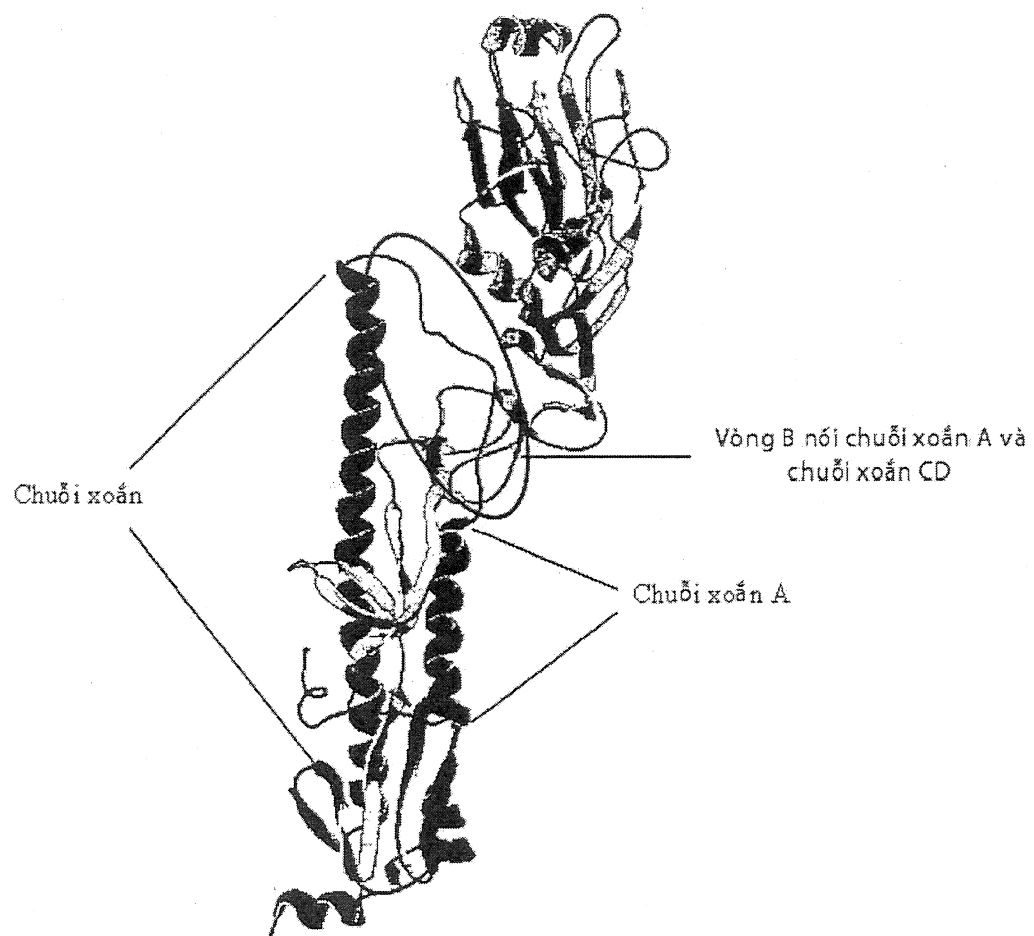
6. Polypeptit theo điểm bất kỳ trong số các điểm nêu trên, trong đó polypeptit này không chứa trình tự tín hiệu.

7. Polypeptit theo điểm bất kỳ trong số các điểm nêu trên, trong đó polypeptit này không chứa trình tự nội bào và xuyên màng của HA.

8. Polypeptit theo điểm 7, trong đó trình tự axit amin từ vị trí 523, 524, 525, 526, 527, 528, 529 hoặc 530 của vùng HA2 đến đầu gốc tận cùng đầu C được loại bỏ.

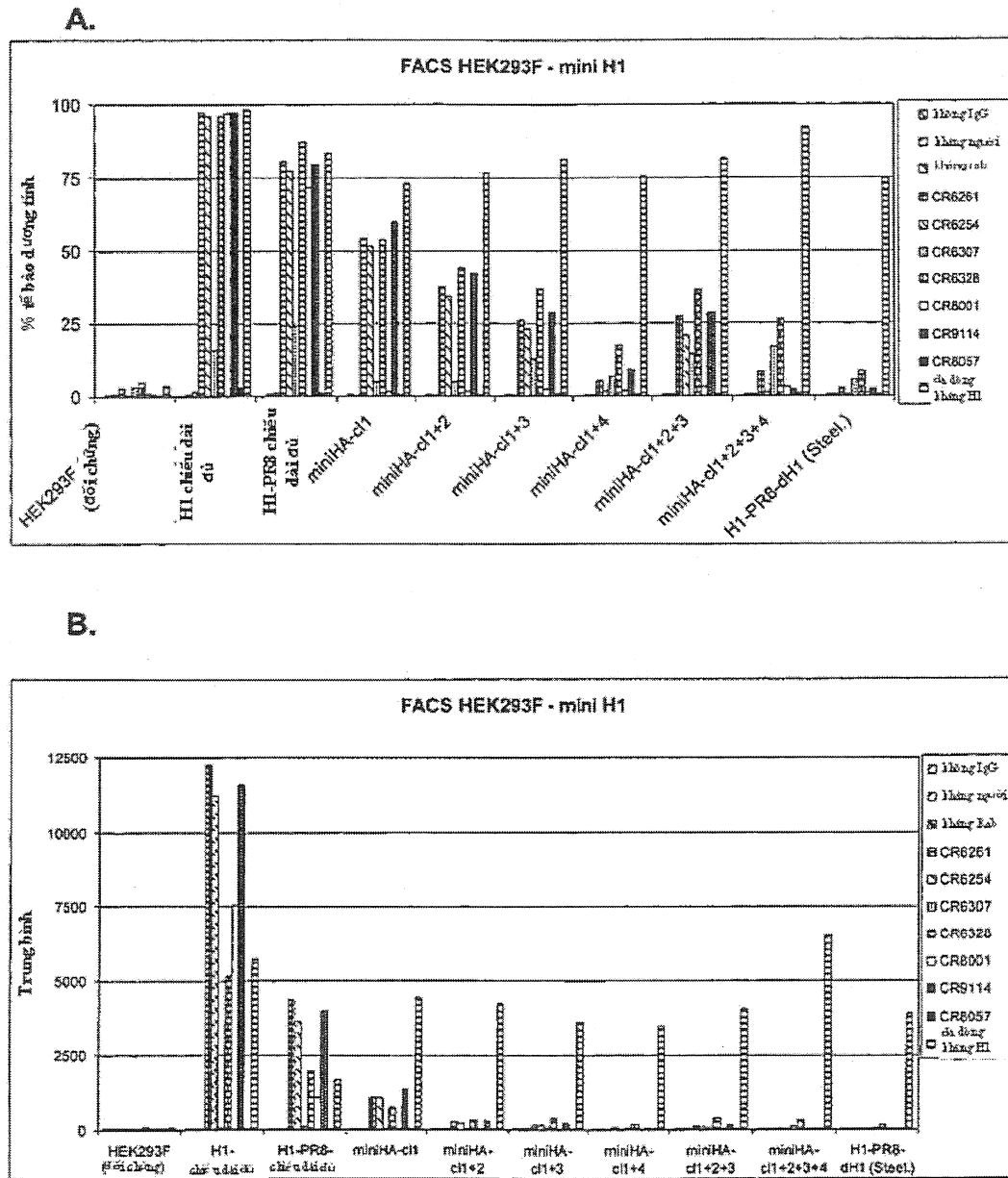
9. Polypeptit theo điểm 8, trong đó trình tự nội bào và xuyên màng được thay thế bởi trình tự axit amin AGRHHHHHHH (SEQ ID NO: 81) hoặc SGRSLVPRGSPGSGYIPEAPRDGQAYVRKDGEWVLLSTFLGHHHHHHH (SEQ ID NO: 82).
10. Polypeptit theo điểm bất kỳ trong số các điểm nêu trên, chứa trình tự axit amin được chọn từ nhóm bao gồm SEQ ID NO: 44-47, 49-53, 144-147, 150-157, và 181-185.
11. Axit nucleic mã hoá polypeptit theo điểm bất kỳ trong số các điểm 1-10.
12. Chế phẩm miễn dịch chứa polypeptit theo điểm bất kỳ trong số các điểm 1-10, và/hoặc phân tử axit nucleic theo điểm 11.

FIG. 1.



2/52

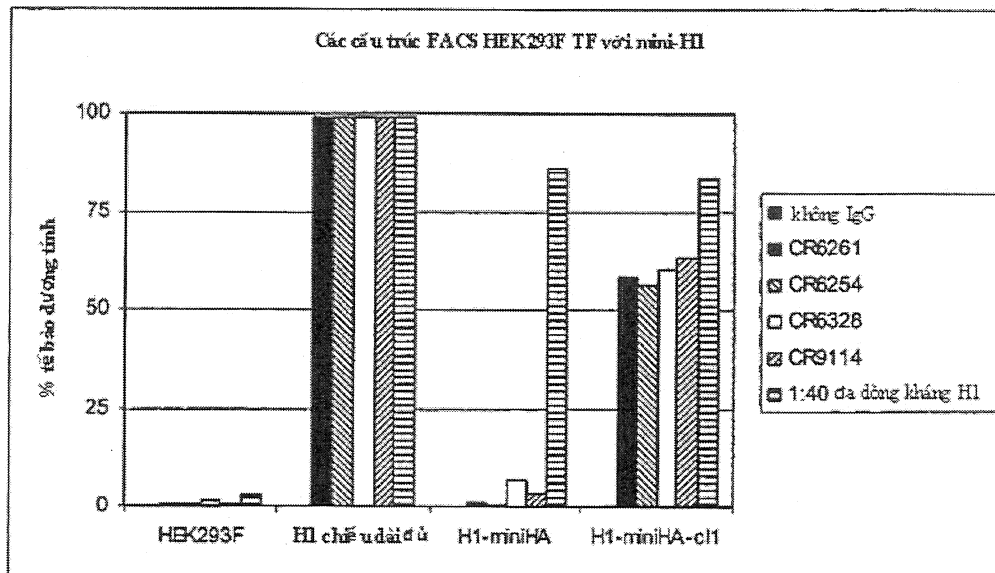
FIG. 2



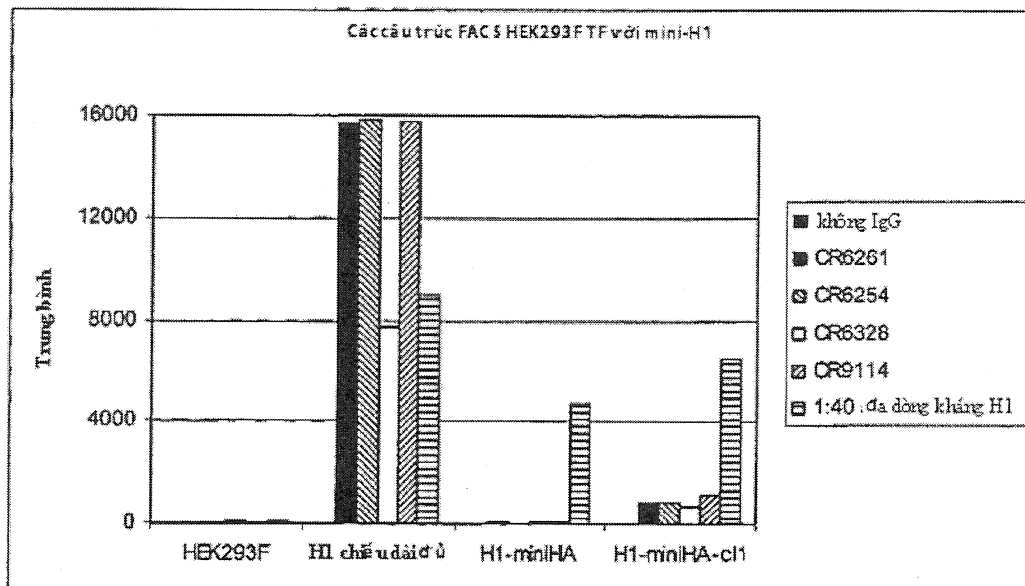
3/52

FIG. 3

A.

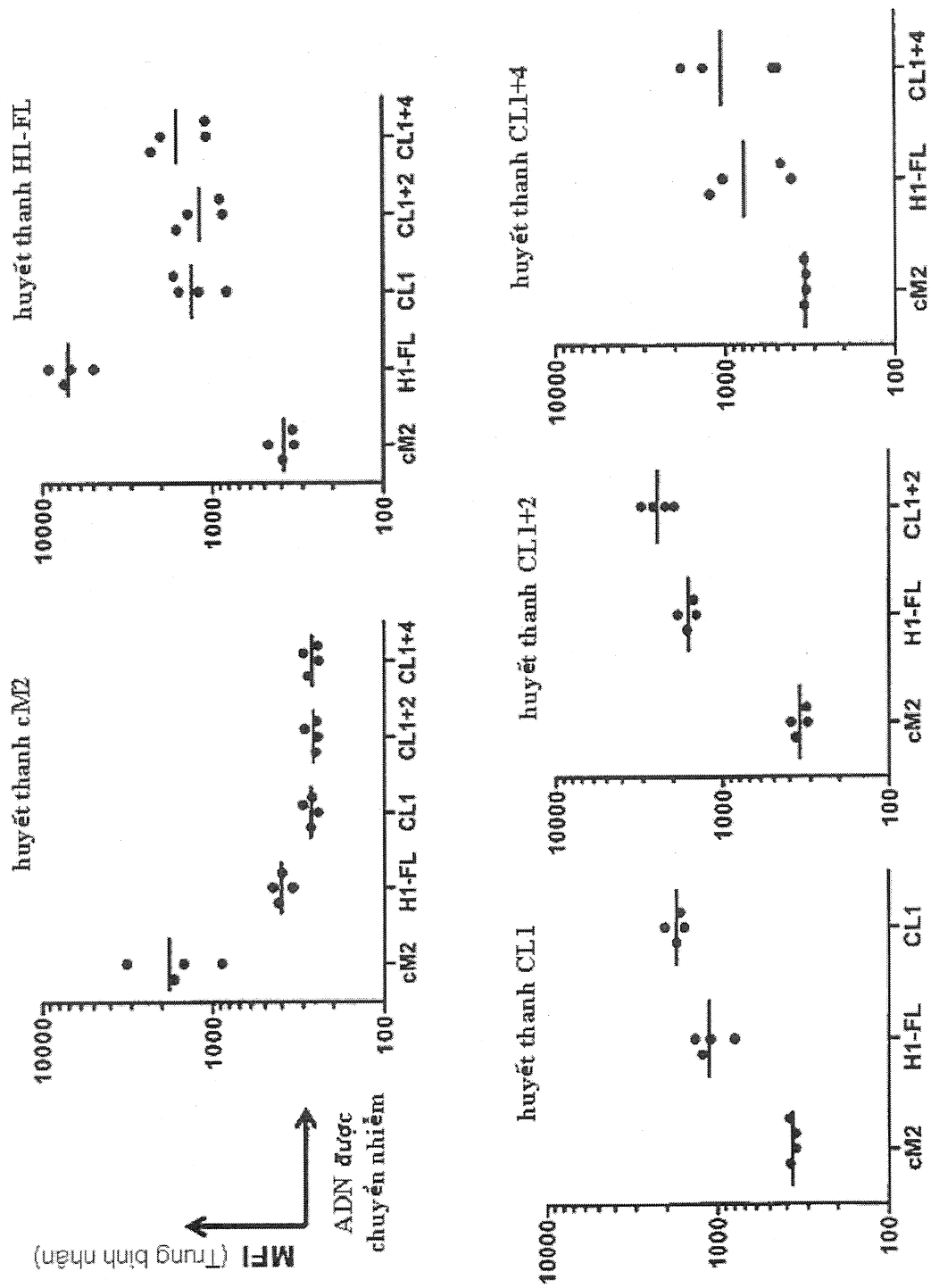


B.



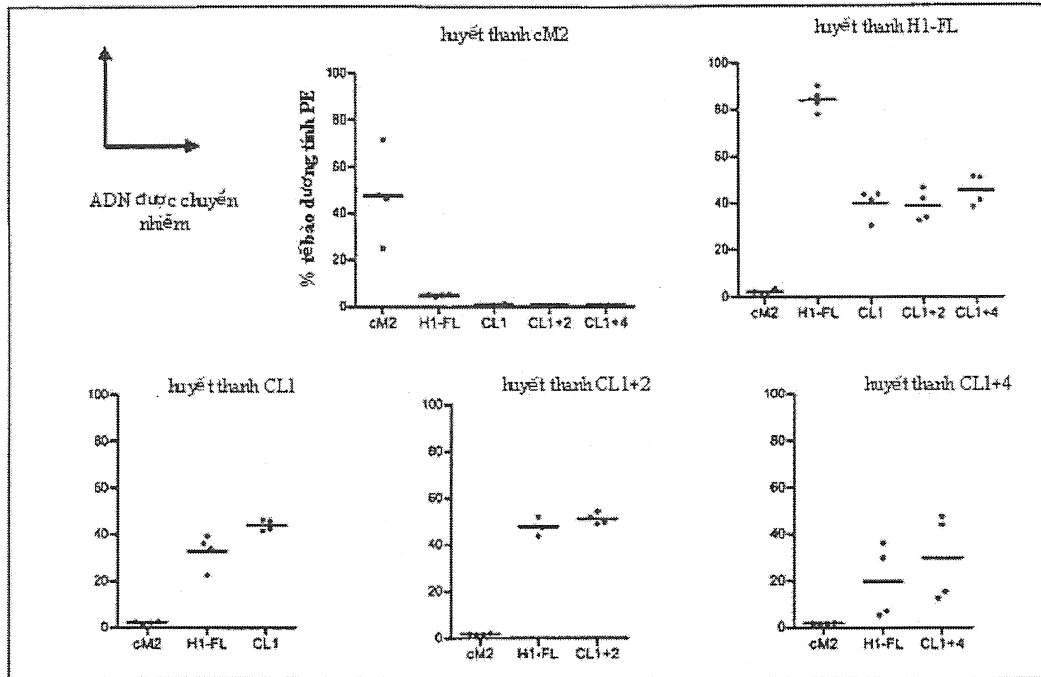
4/52

Fig 4A



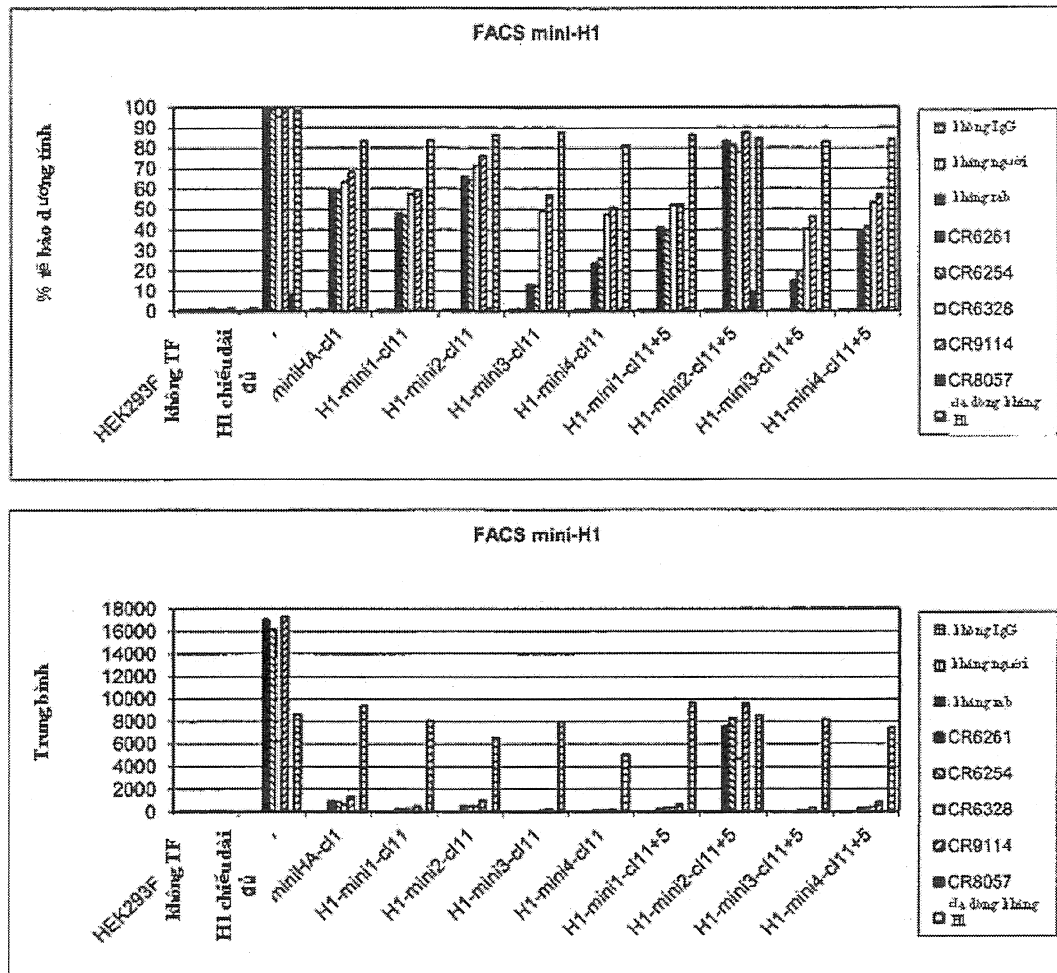
5/52

FIG. 4B



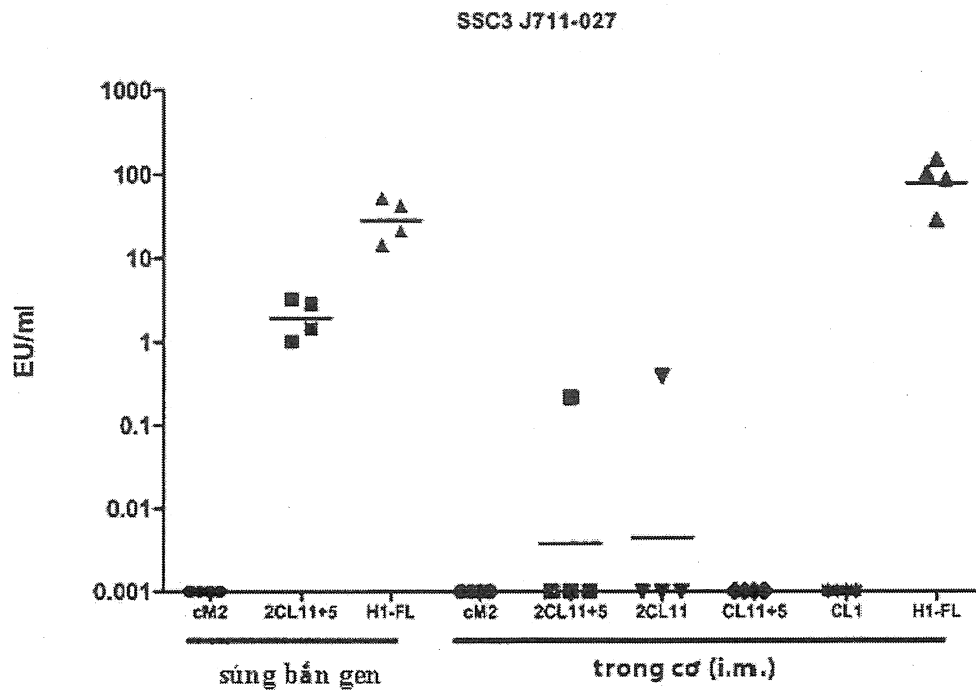
6/52

FIG. 5



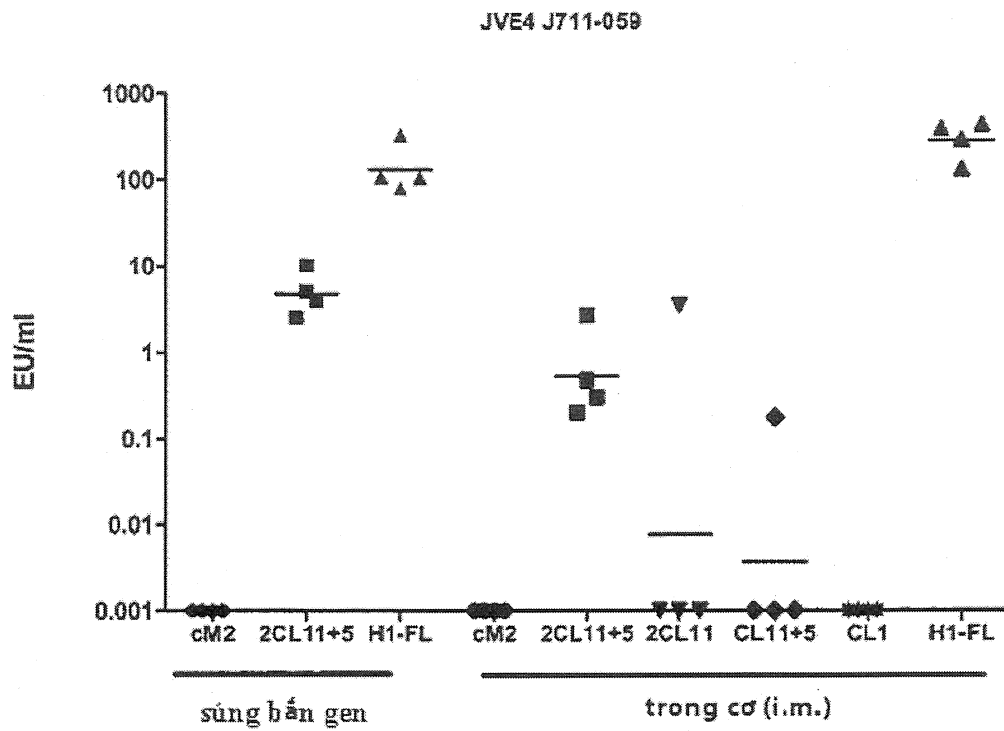
7/52

FIG. 6A



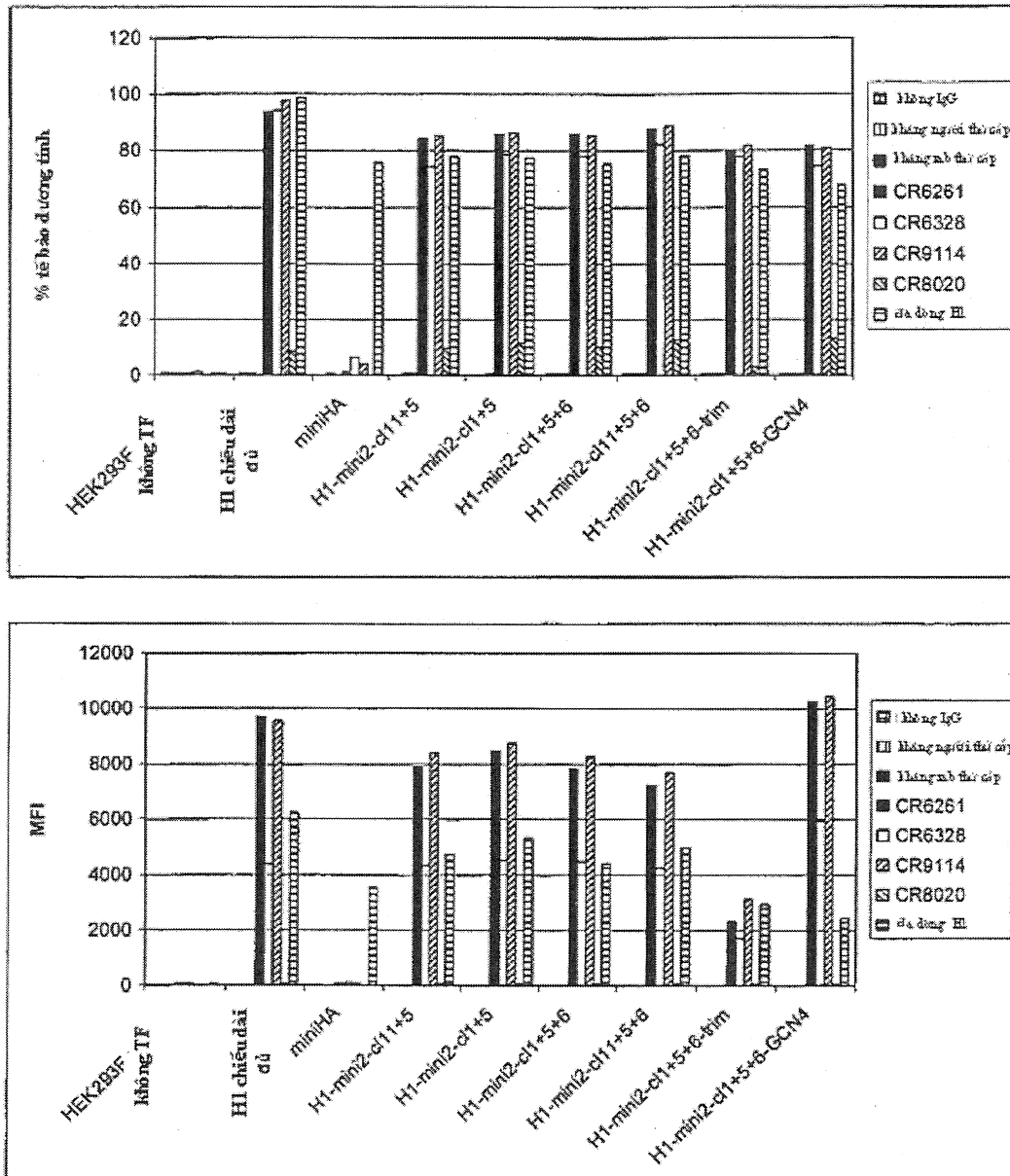
8/52

FIG. 6B



9/52

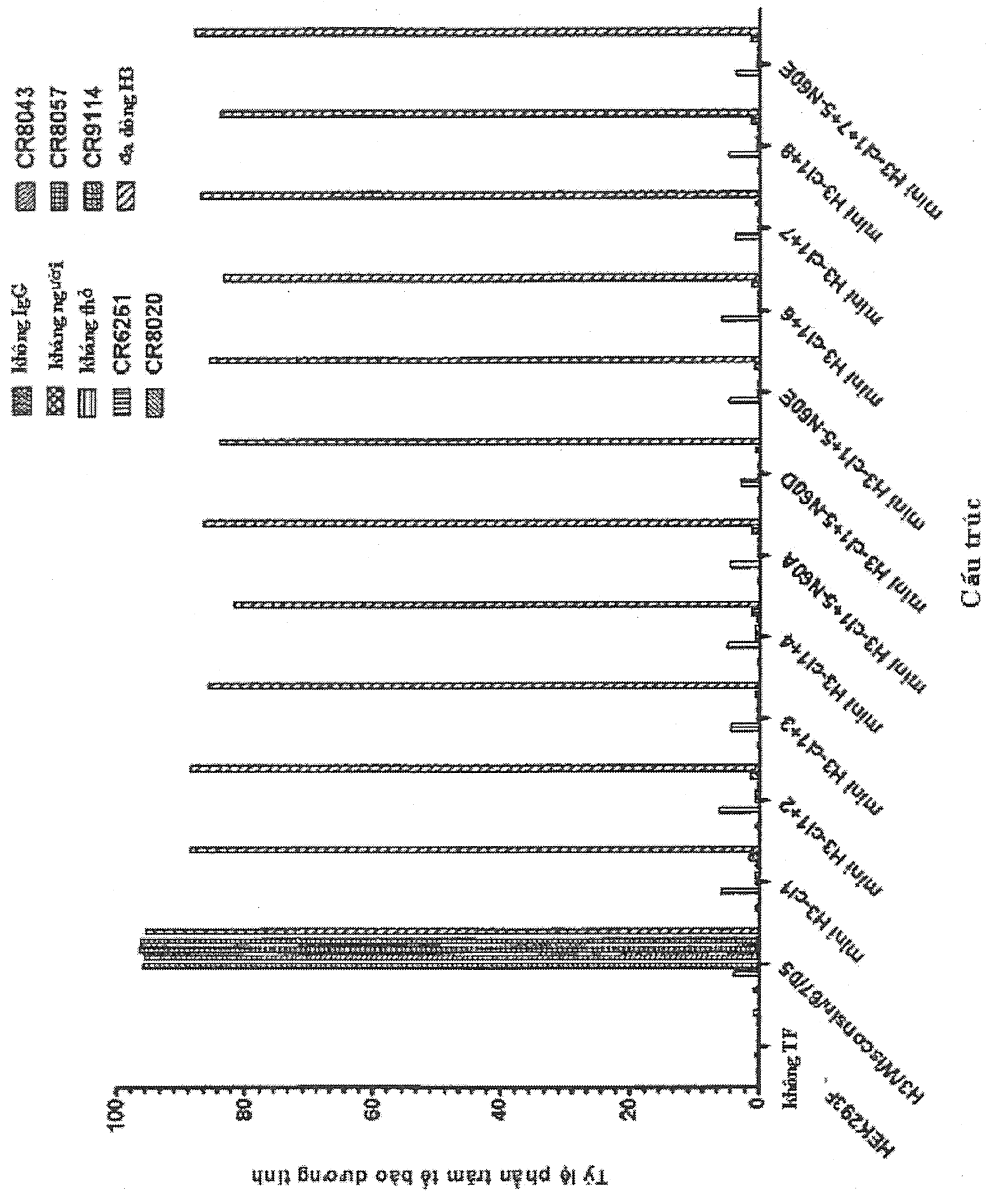
FIG. 7



10/52

FIG. 8

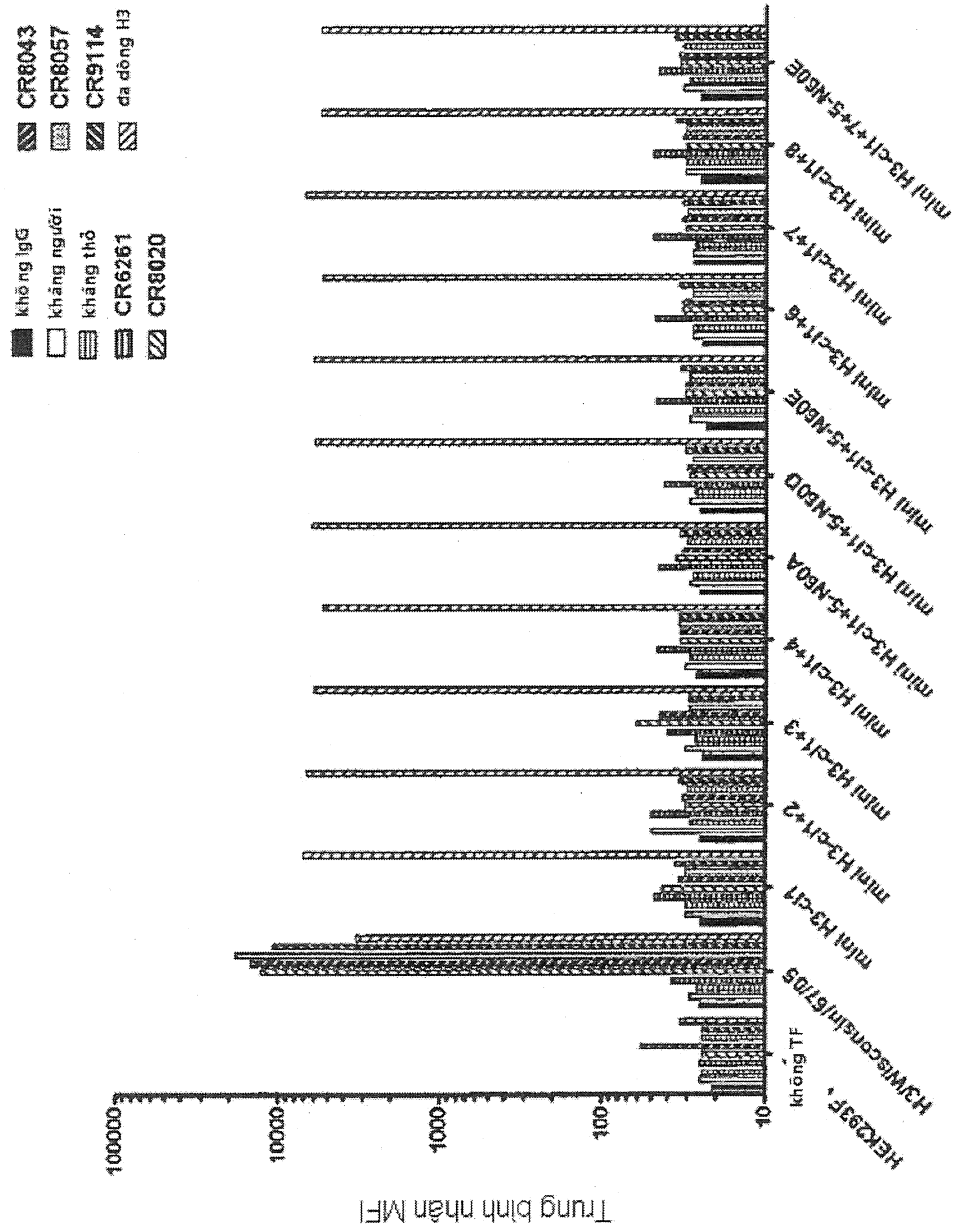
A.



11/52

FIG. 8

B.



12/52

FIG. 9A

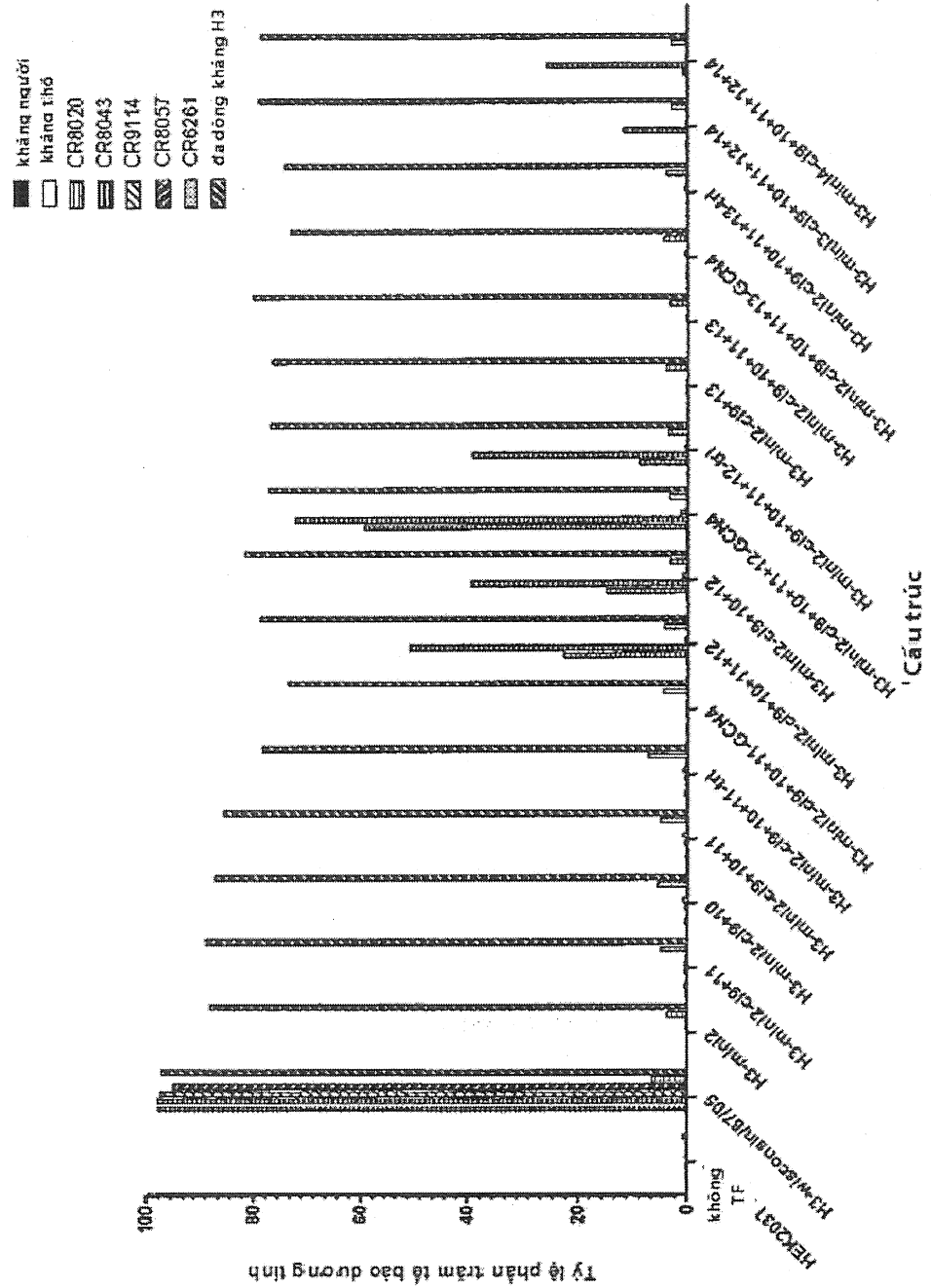
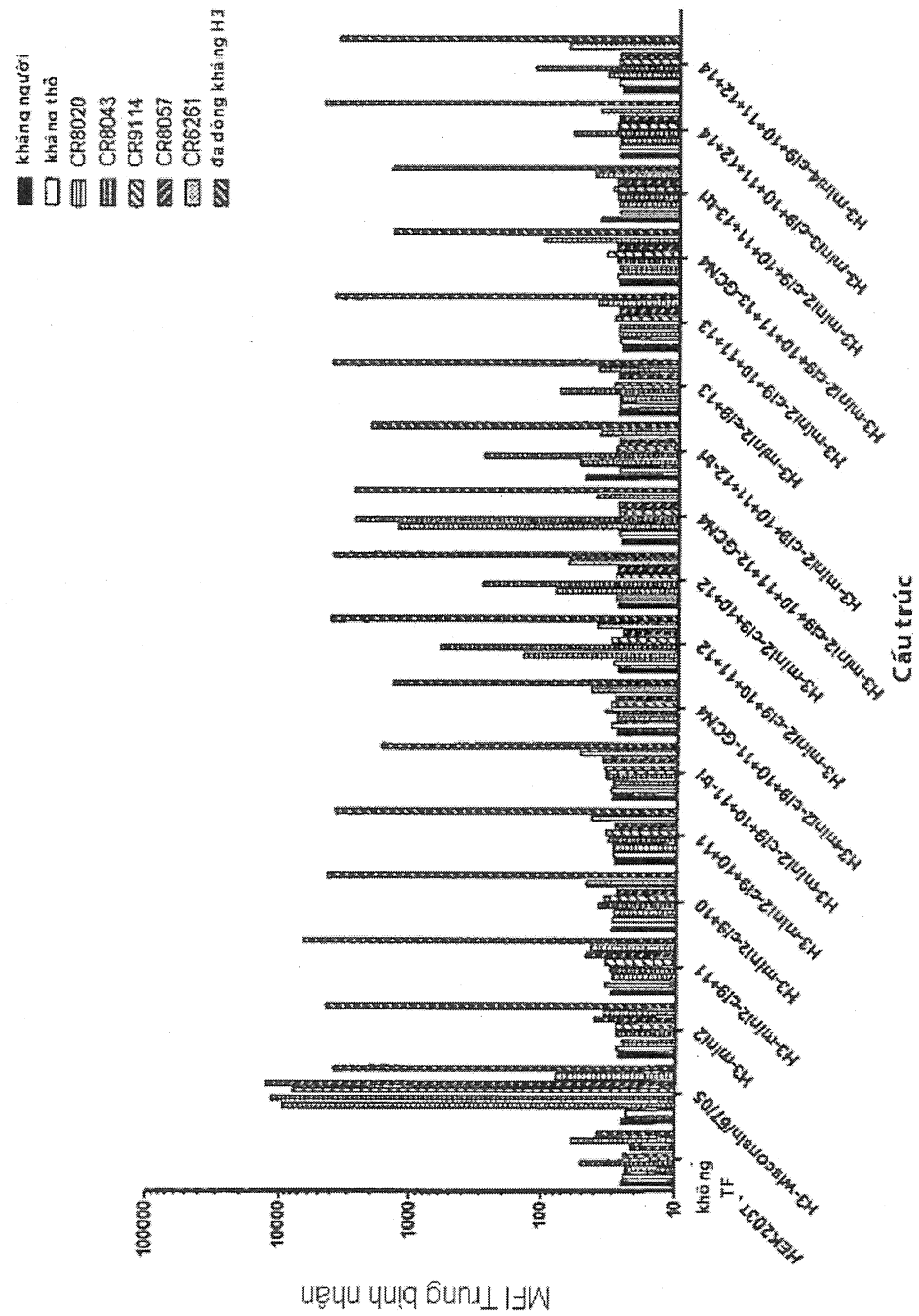


FIG. 9B



14/52

FIG. 10A

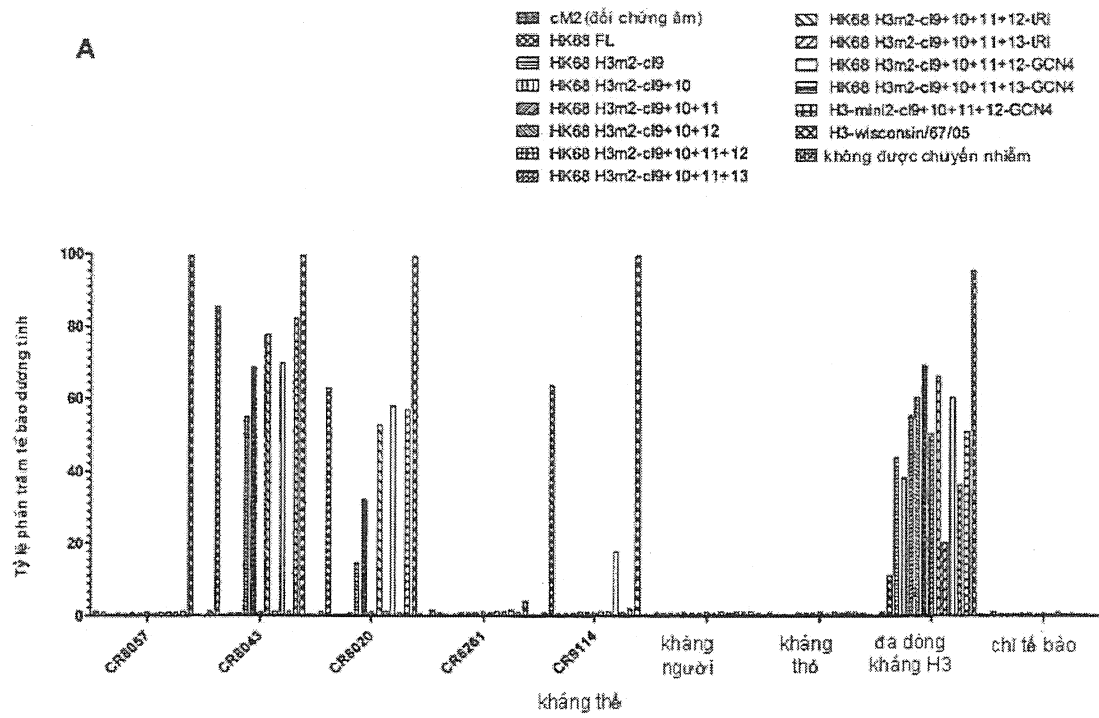
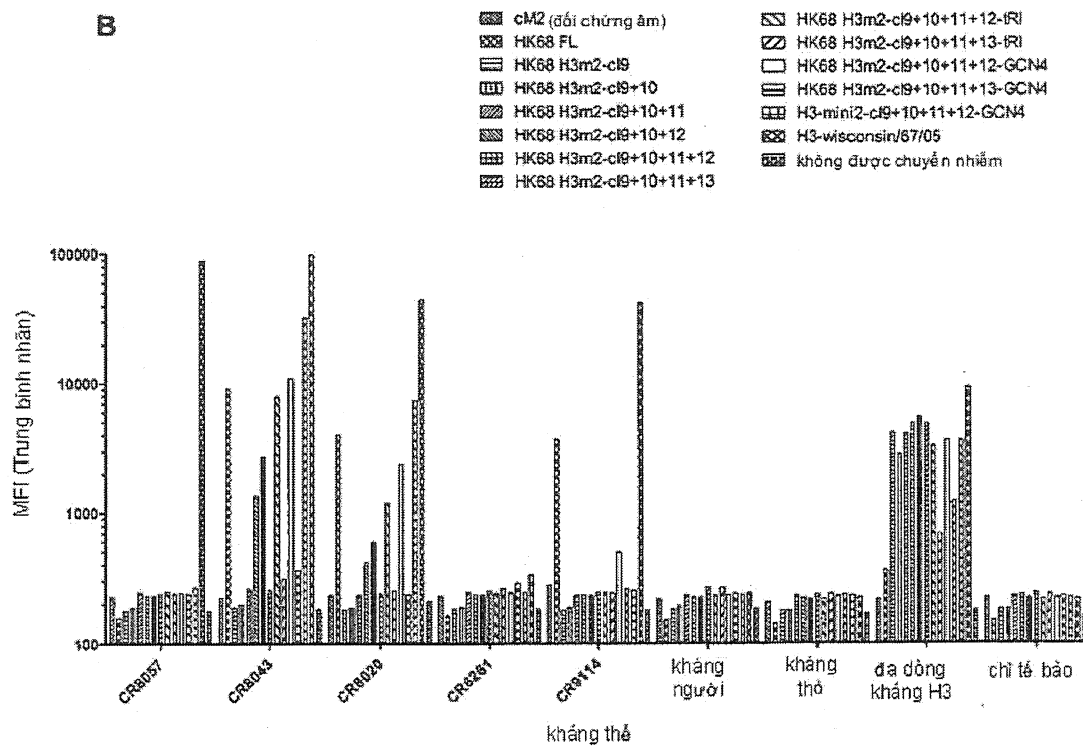
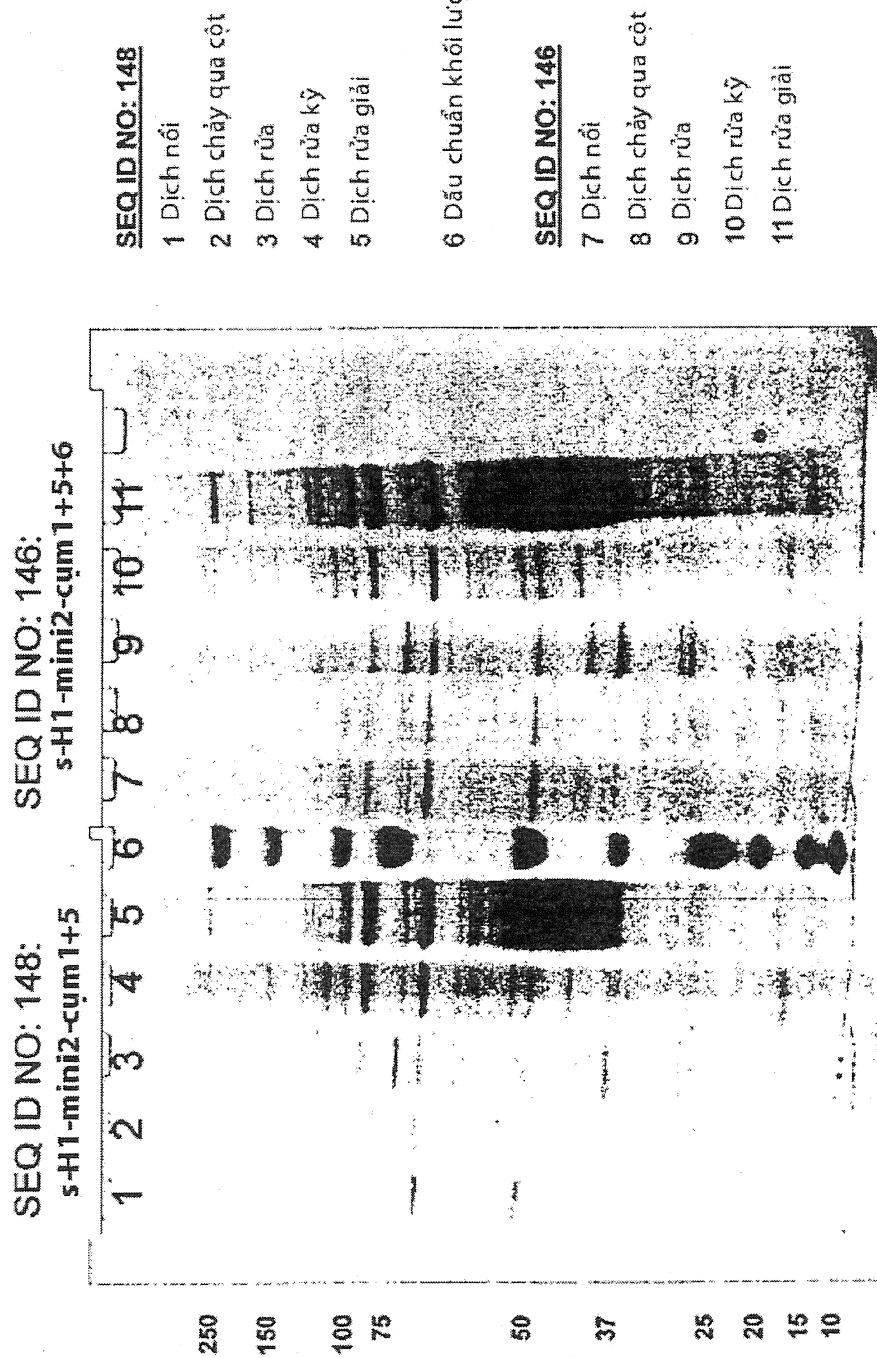


FIG. 10B



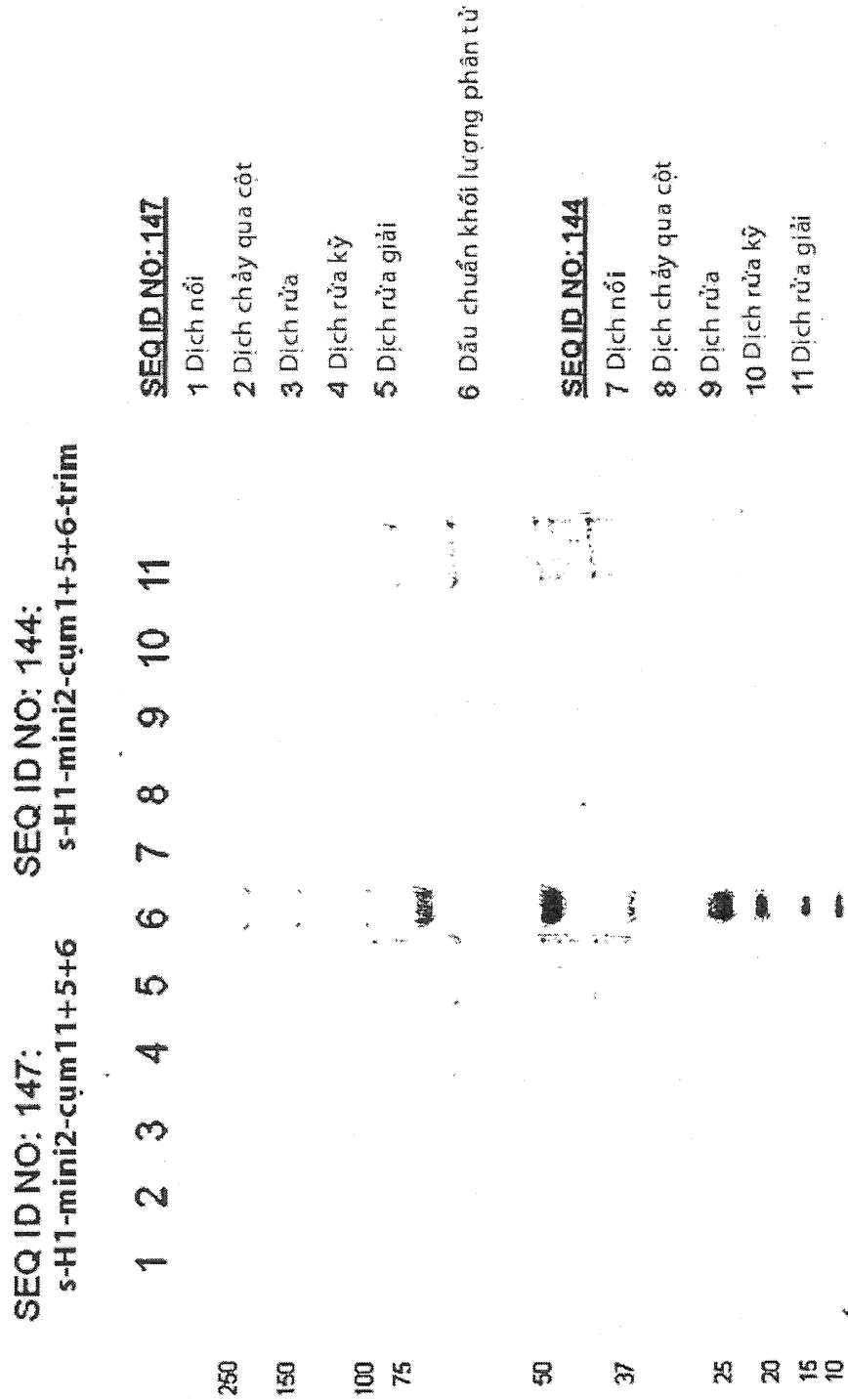
16/52

FIG. 11A



17/52

FIG. 11B

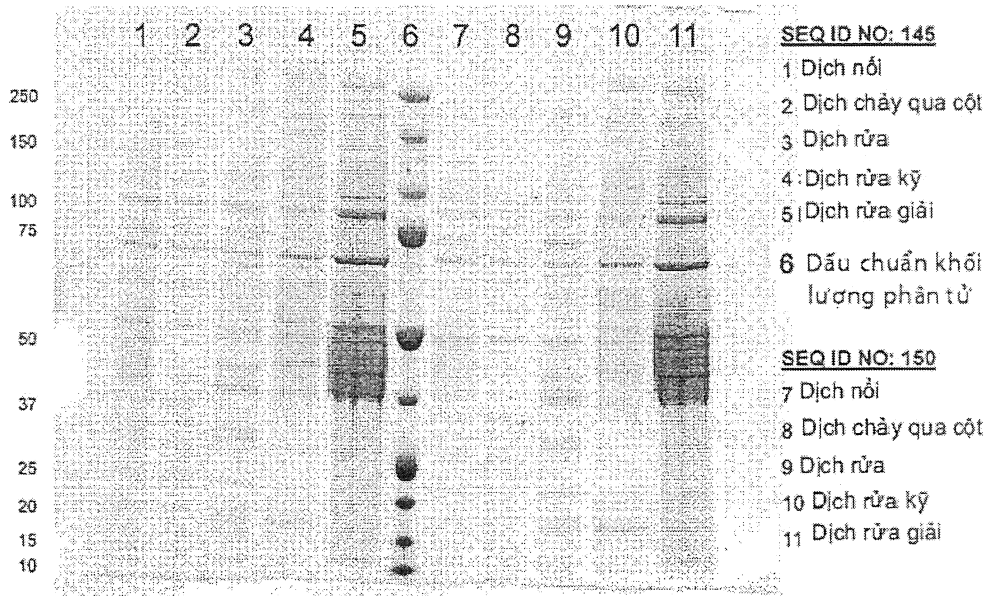


18/52

FIG. 11C

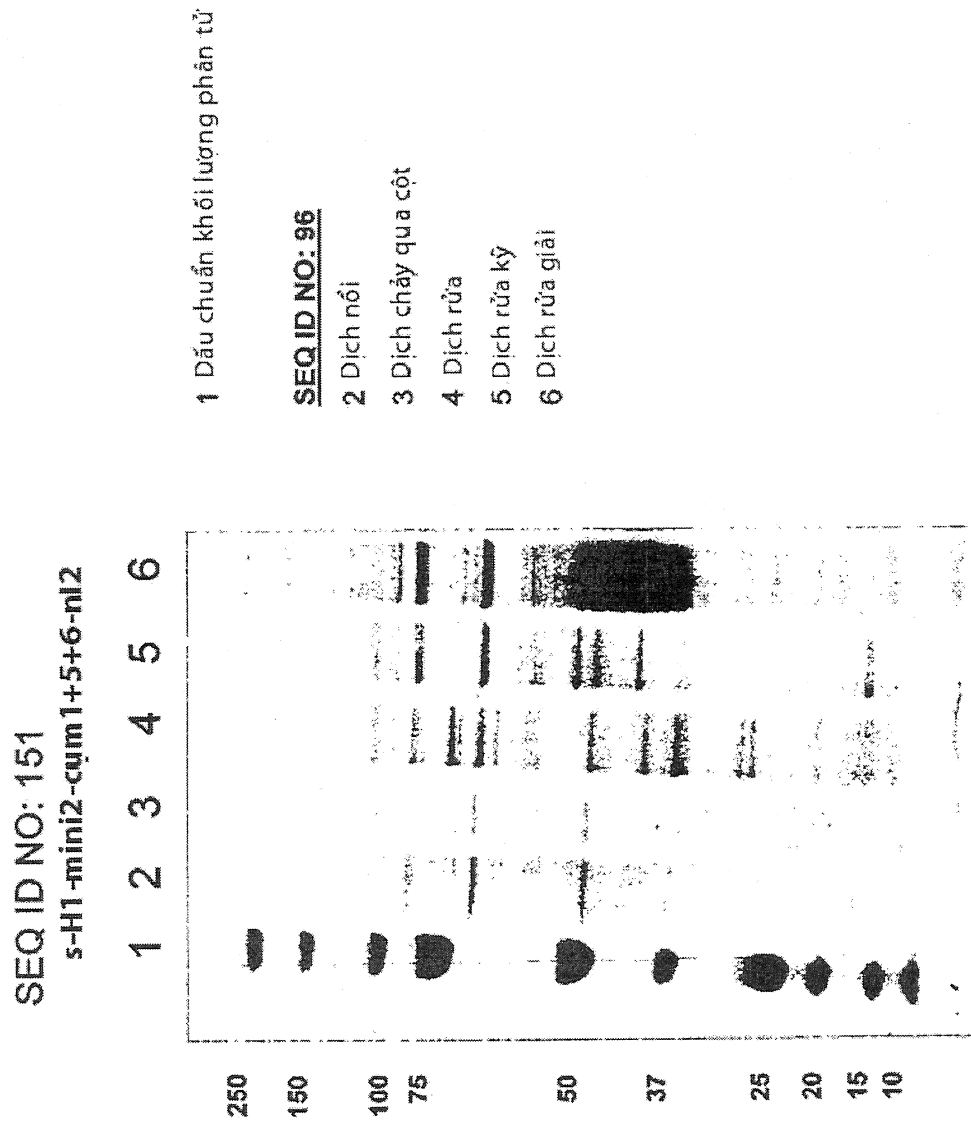
SEQ ID NO: 145:
s-H1-mini2-cụm1+5+6-GCN4

SEQ ID NO: 150:
s-H1-mini2-cụm1+5+6-nl



19/52

FIG. 11D



20/52

FIG. 11E

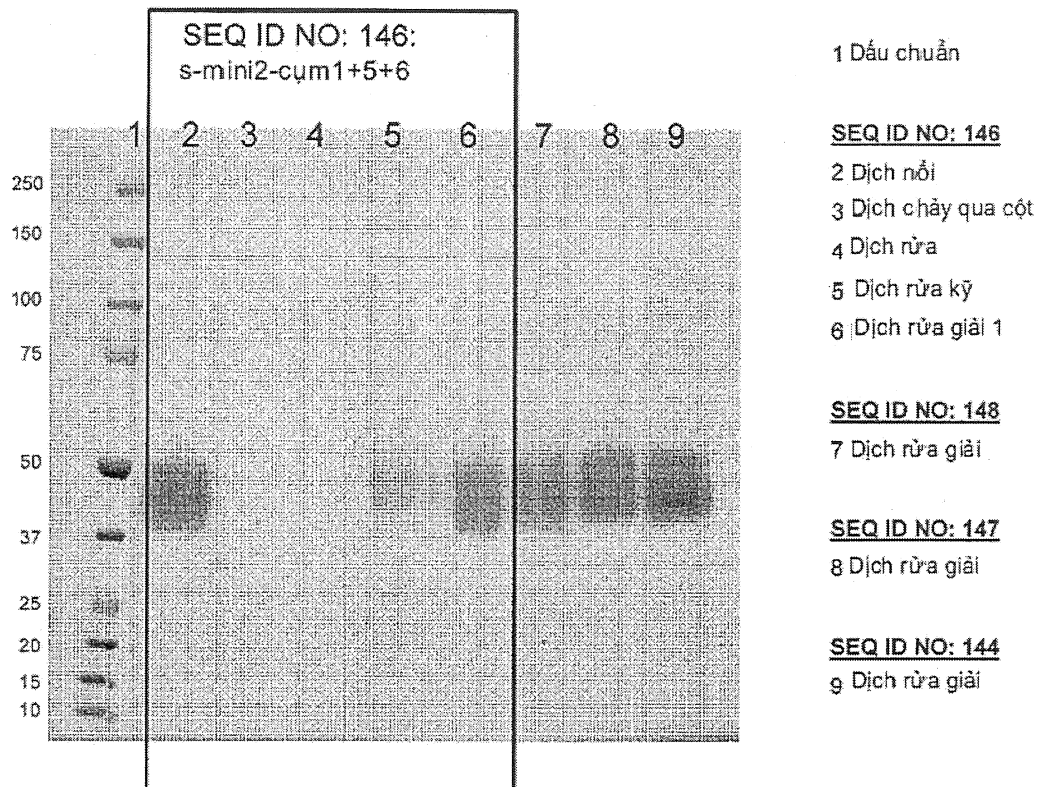
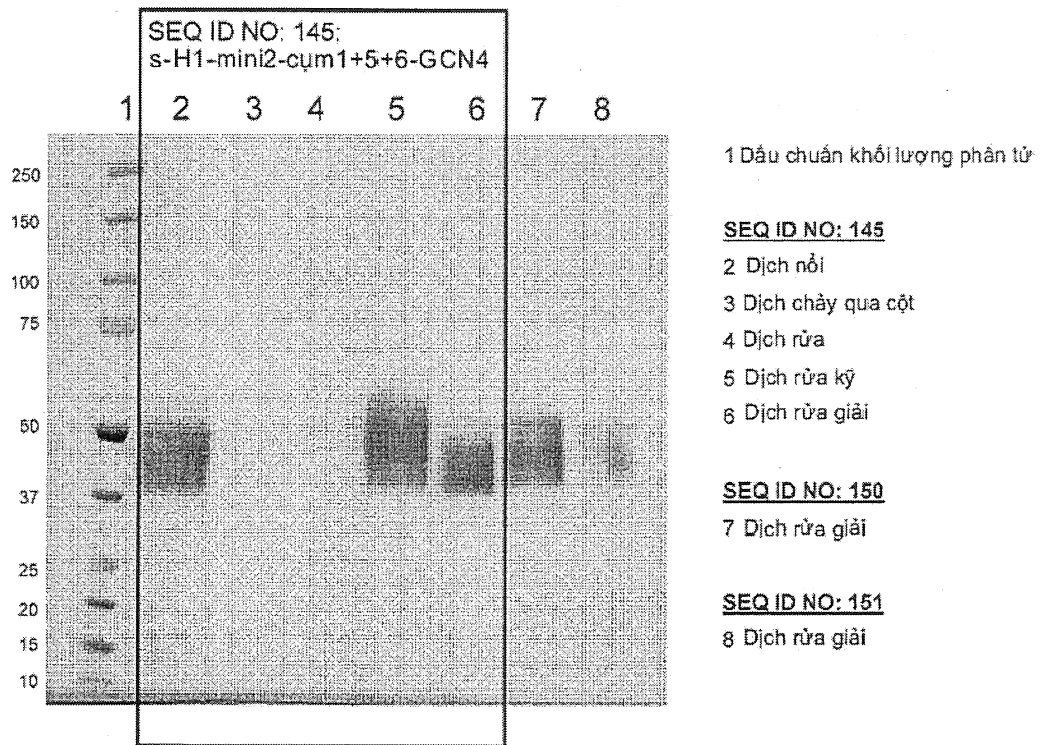


FIG. 11F



22/52

FIG. 12A

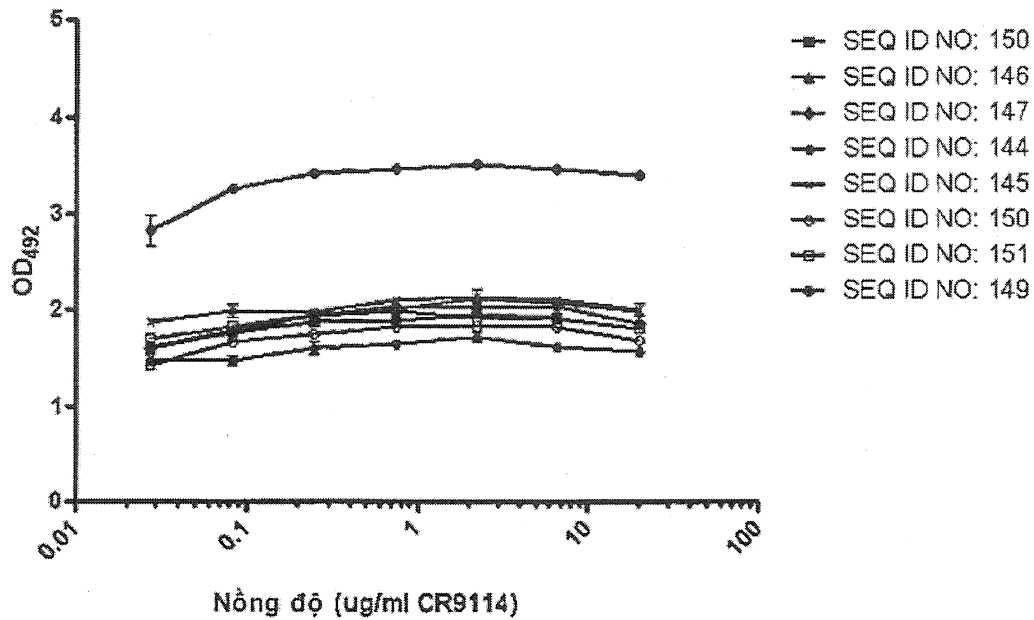
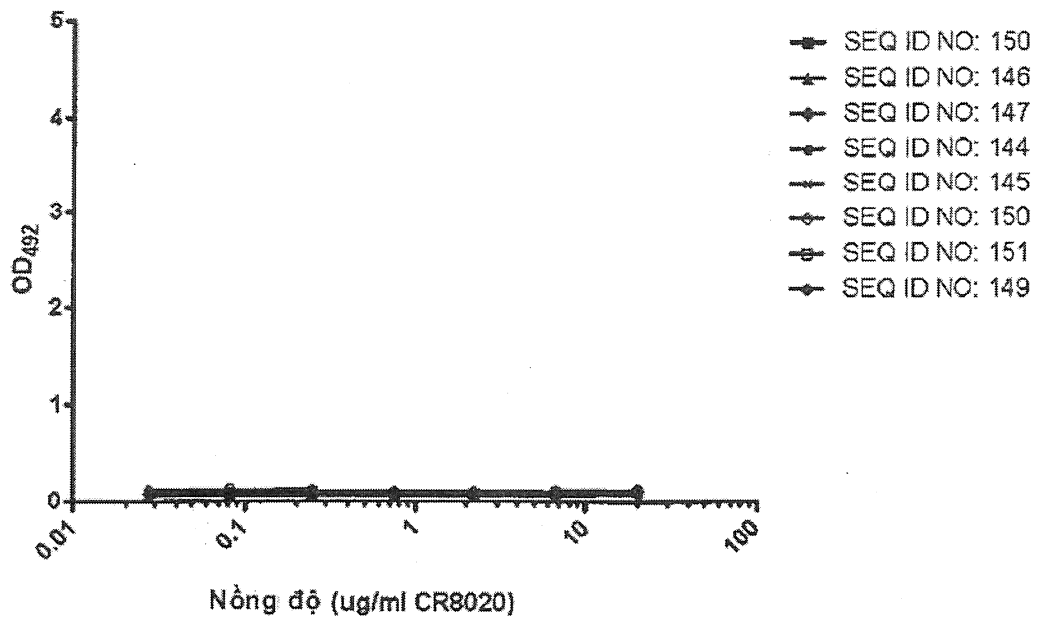
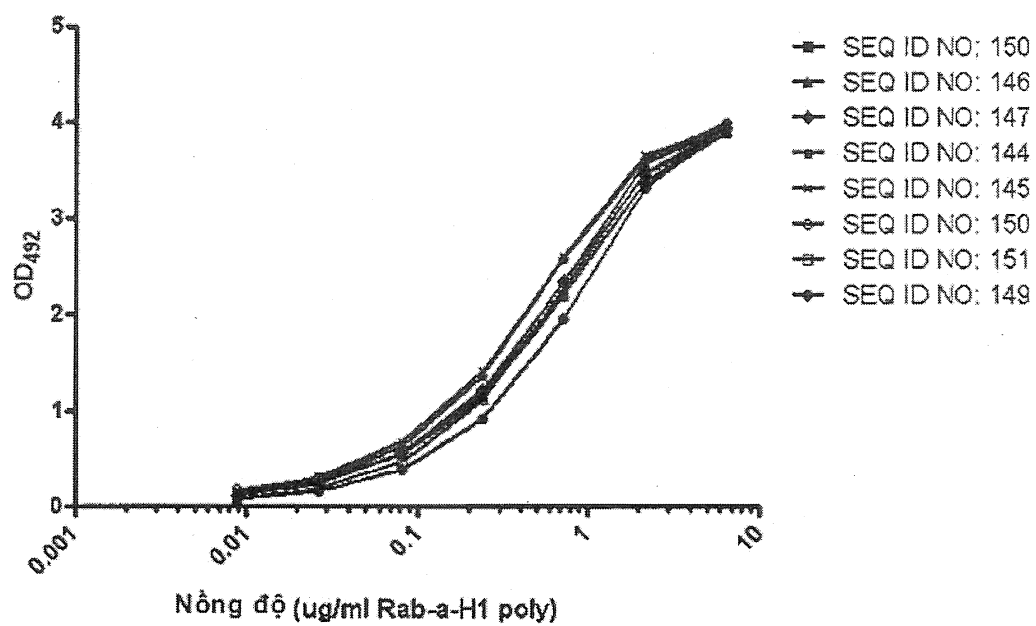


FIG. 12B



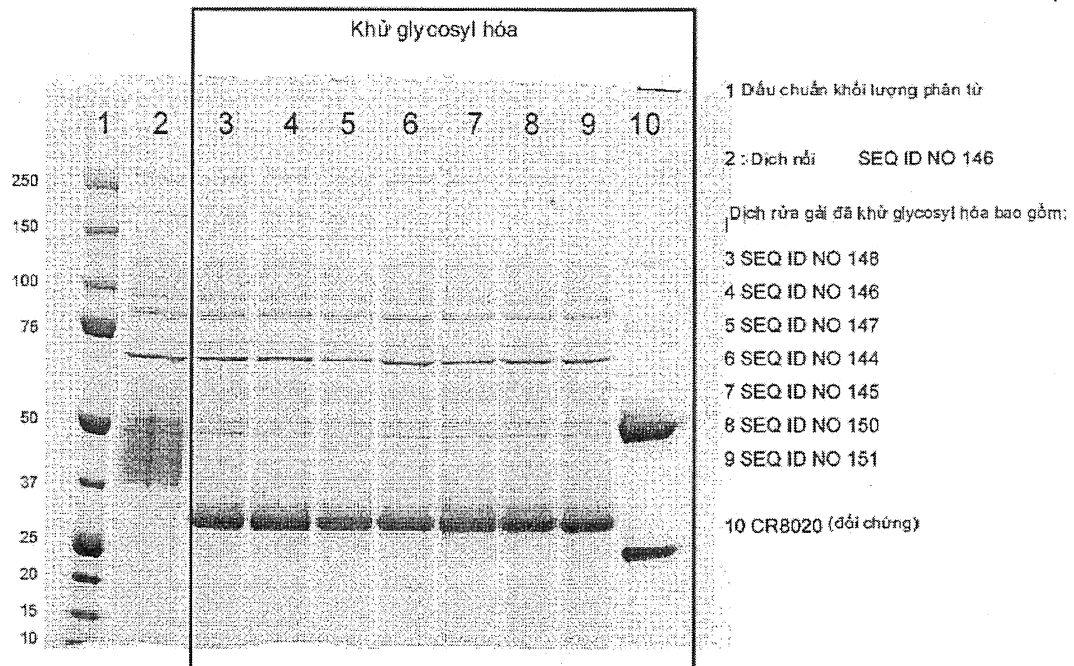
23/52

FIG. 12C



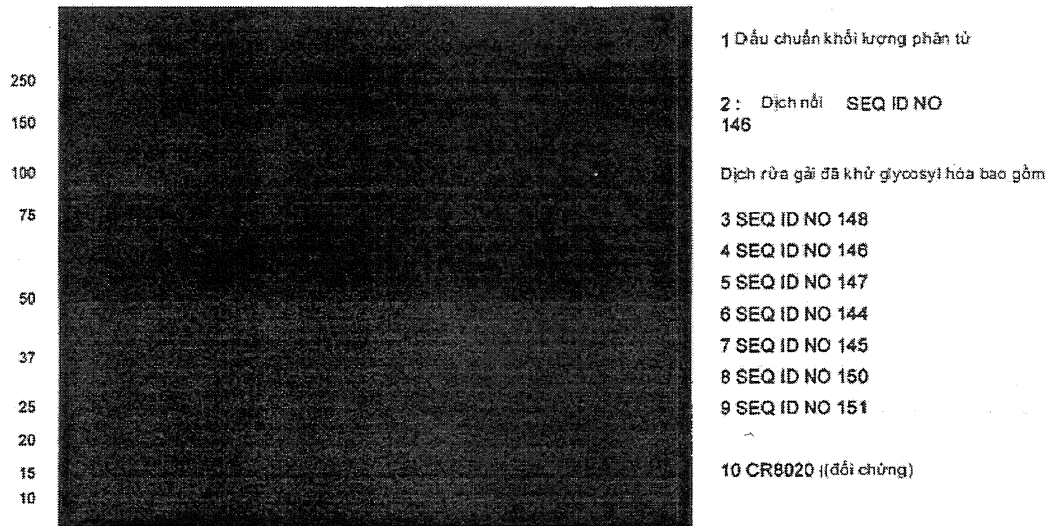
24/52

FIG. 13A



25/52

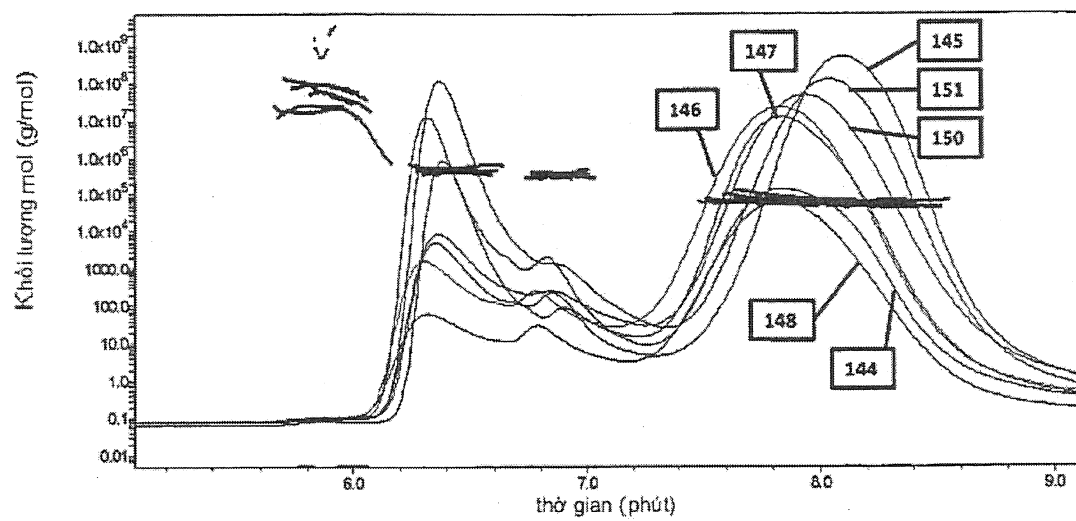
FIG. 13B



Phát hiện: kháng thể đa dòng của thử kháng H1

26/52

FIG. 14



27/52

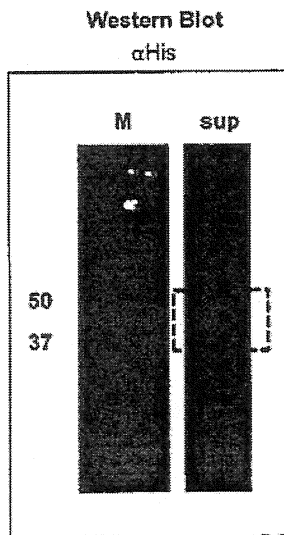


Fig. 15a

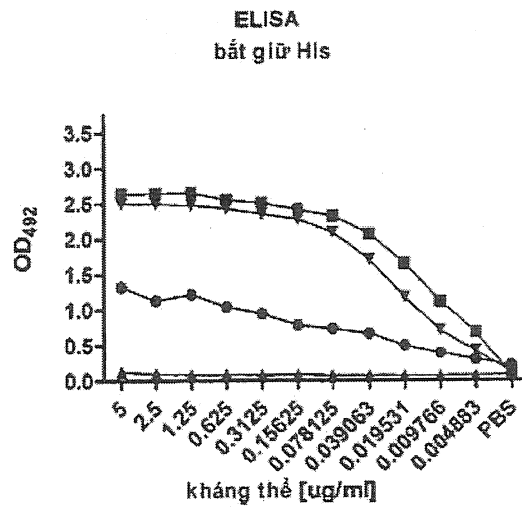
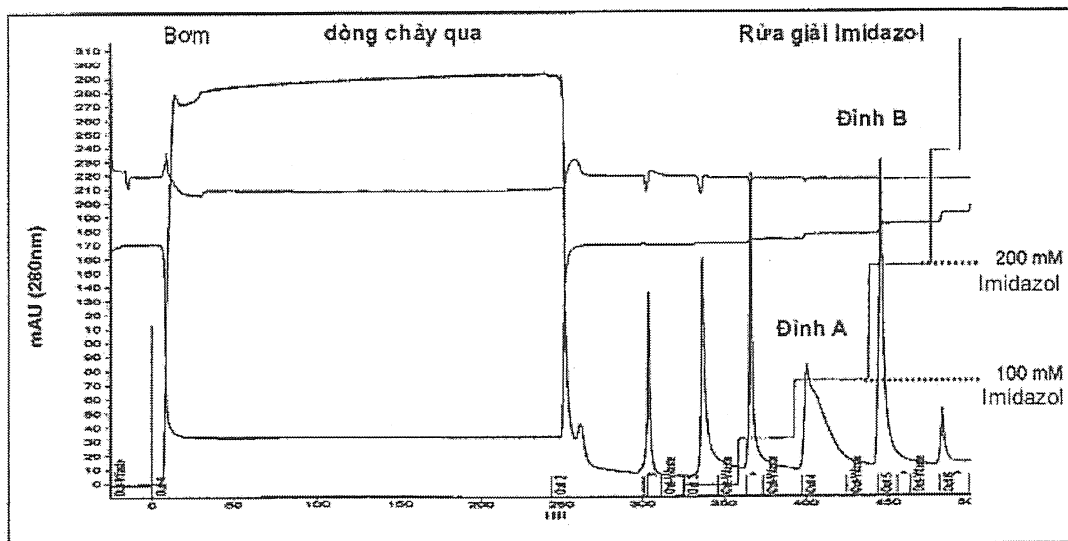


Fig.15b

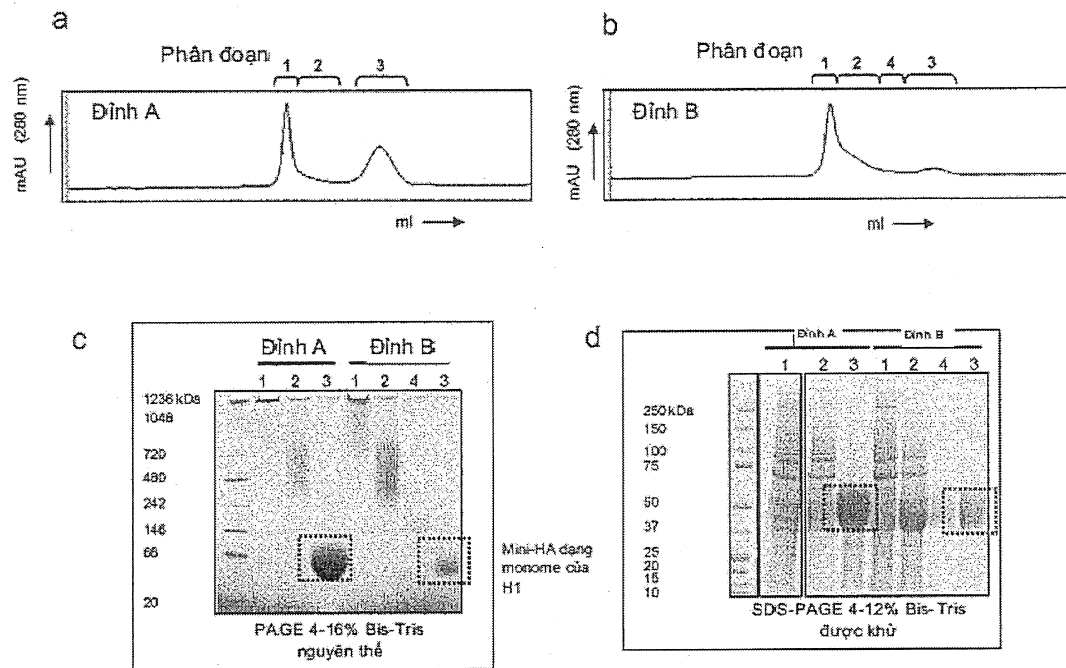
FIG. 16

Nguyên liệu ban đầu; 250ml dịch nổi
Cột: bắt giữ His 5mL
Rửa giải; gradient theo bậc Imidazol



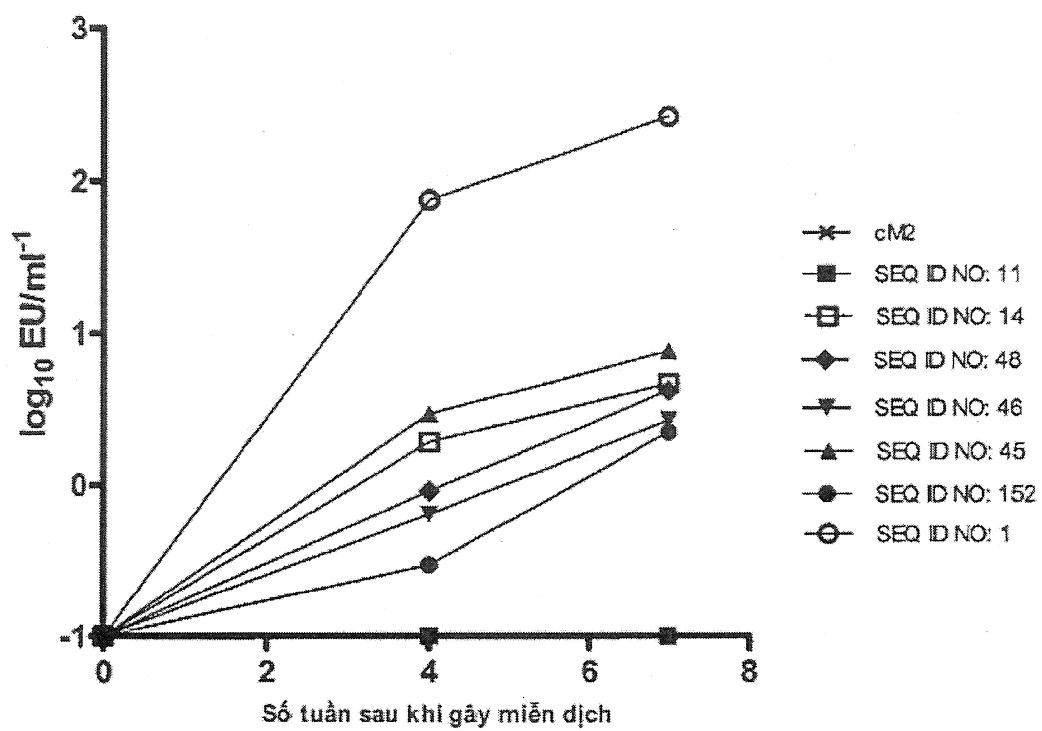
28/52

FIG. 17



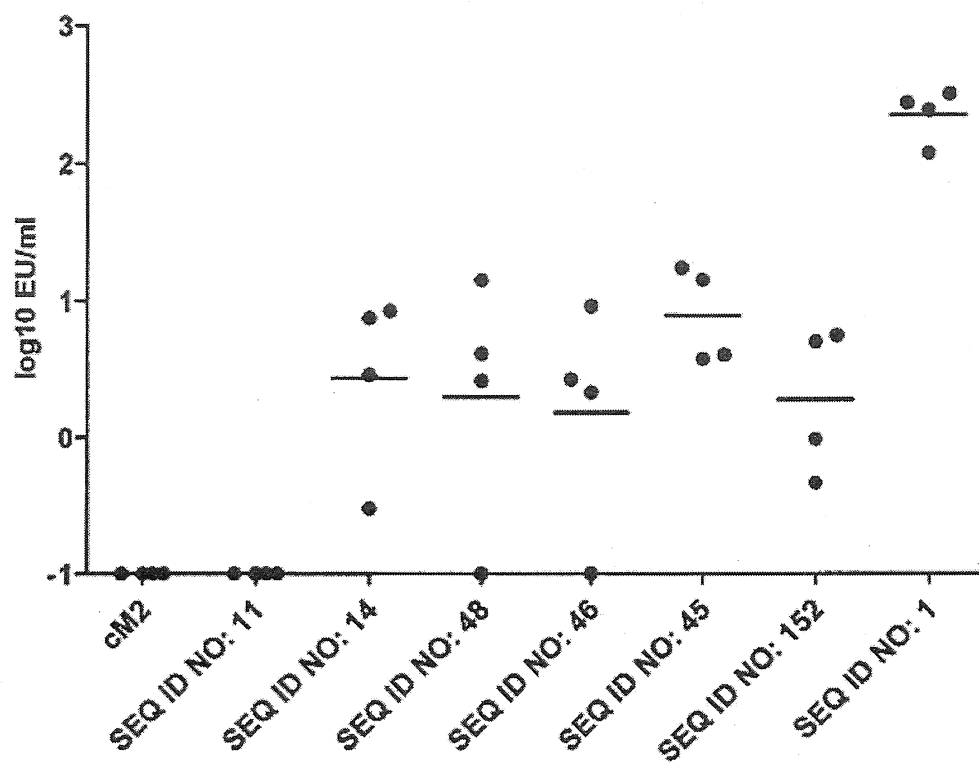
29/52

FIG. 18



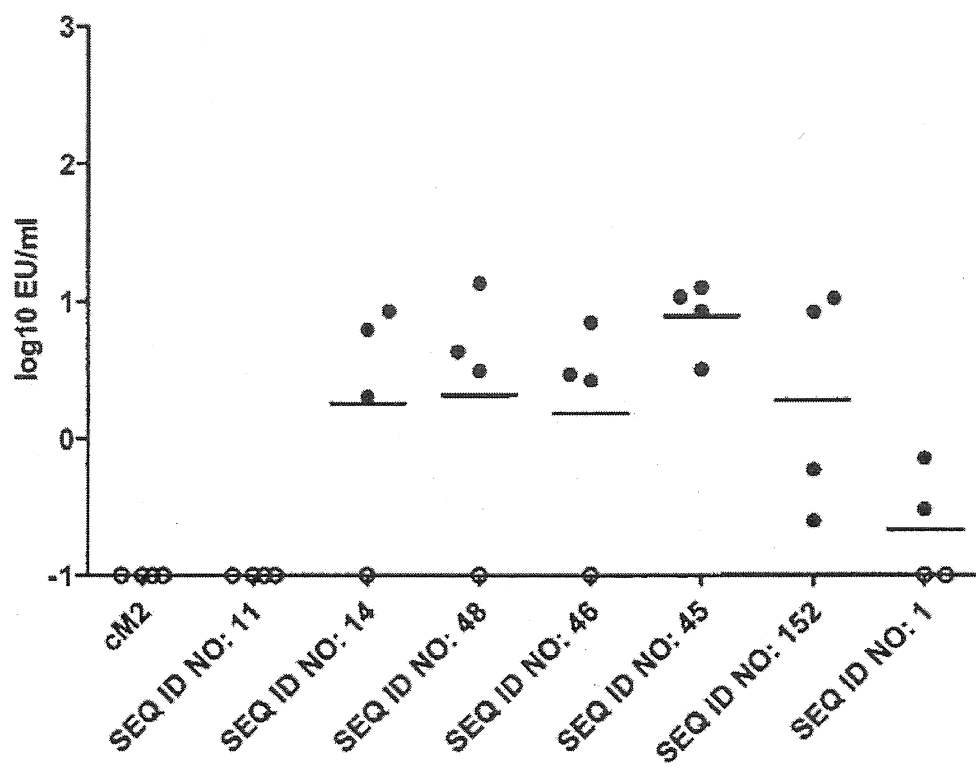
30/52

FIG. 19A



31/52

FIG. 19B



32/52

FIG. 20A

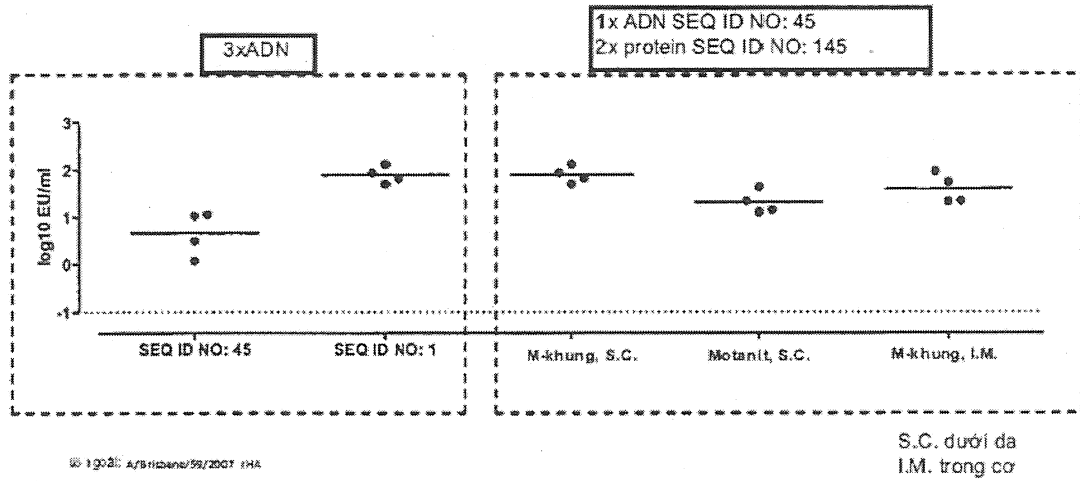
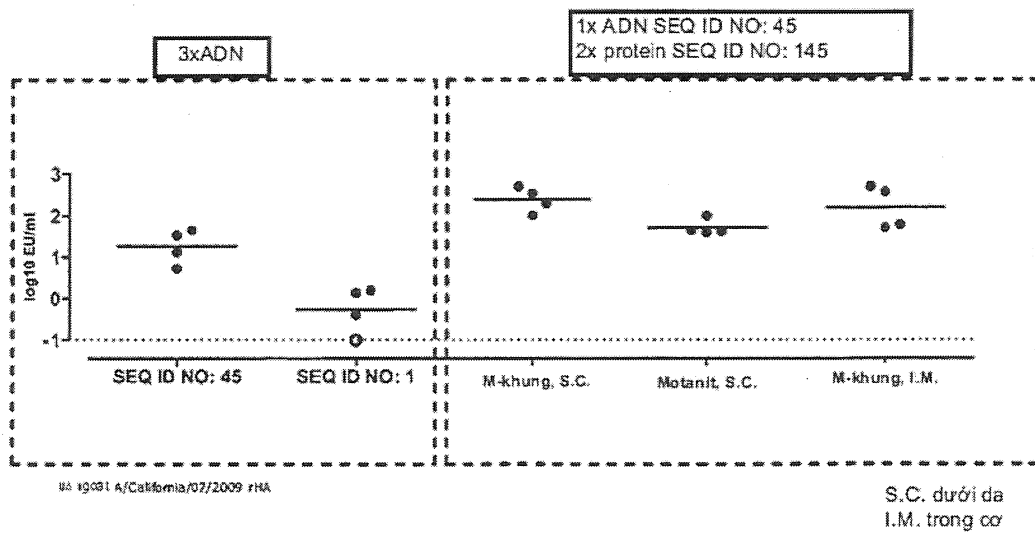


FIG. 20B



33/52

FIG. 20C

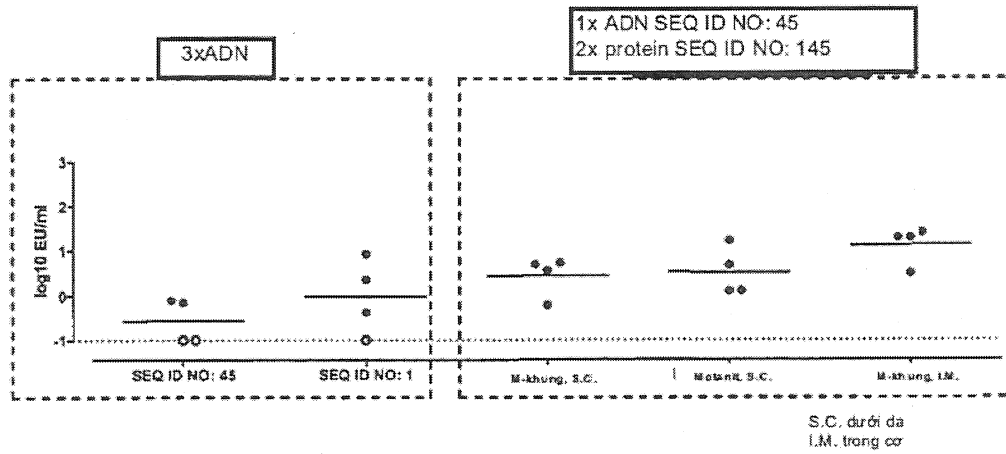


FIG. 20D

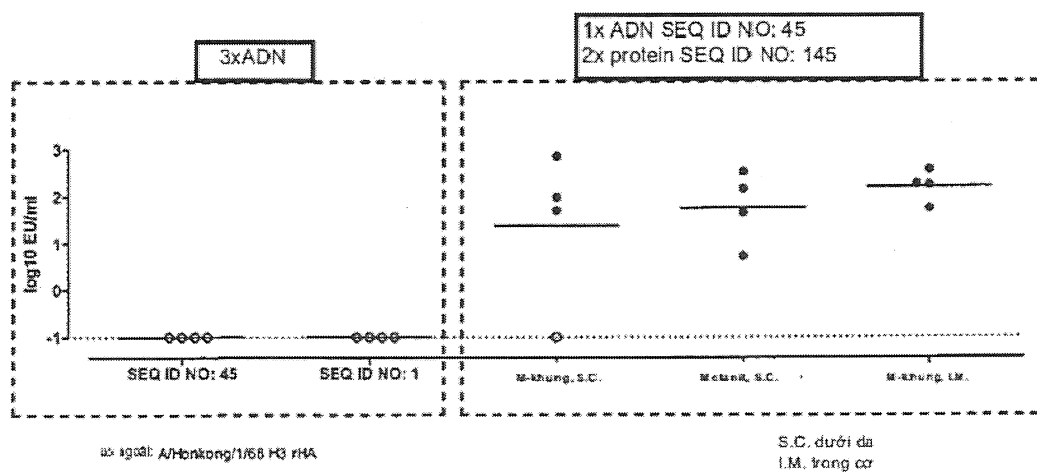
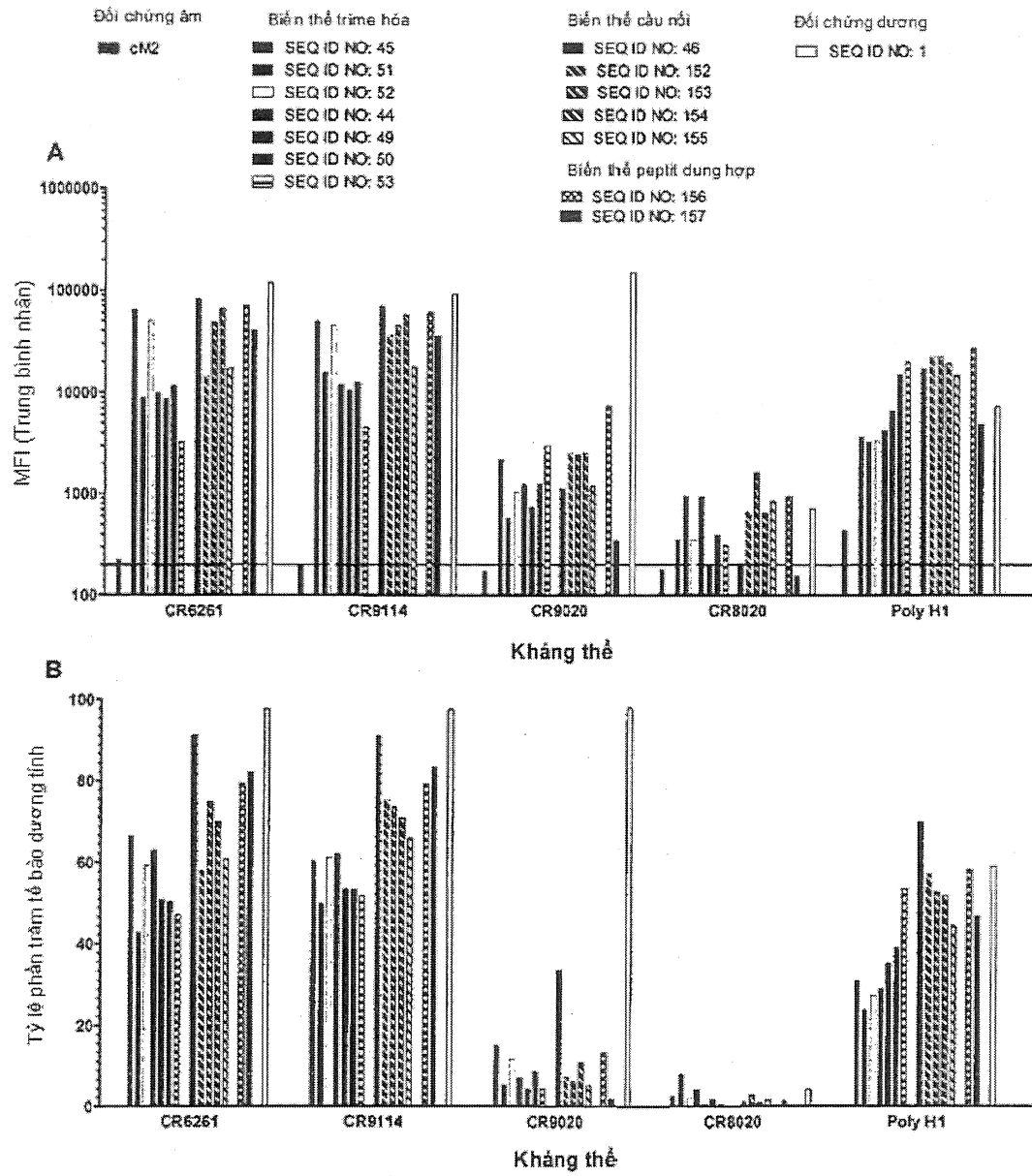


FIG. 21



35/52

FIG. 22A

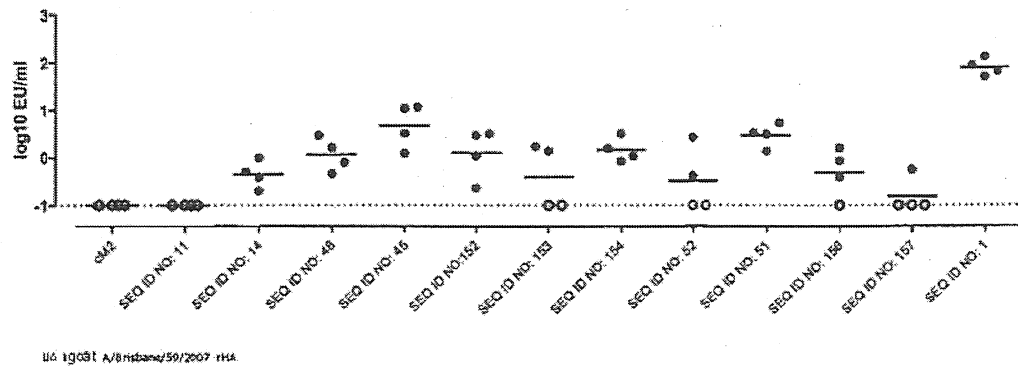
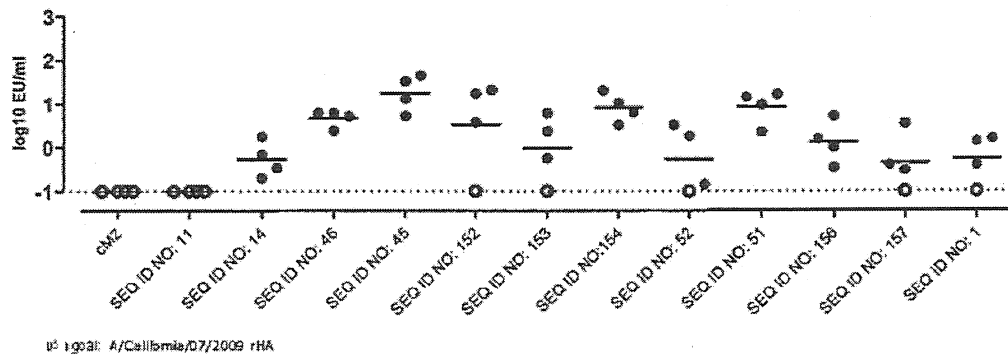
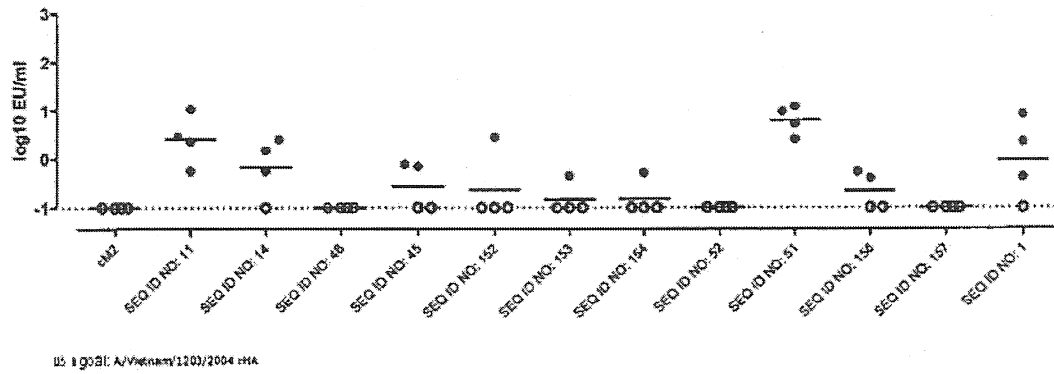


FIG. 22B



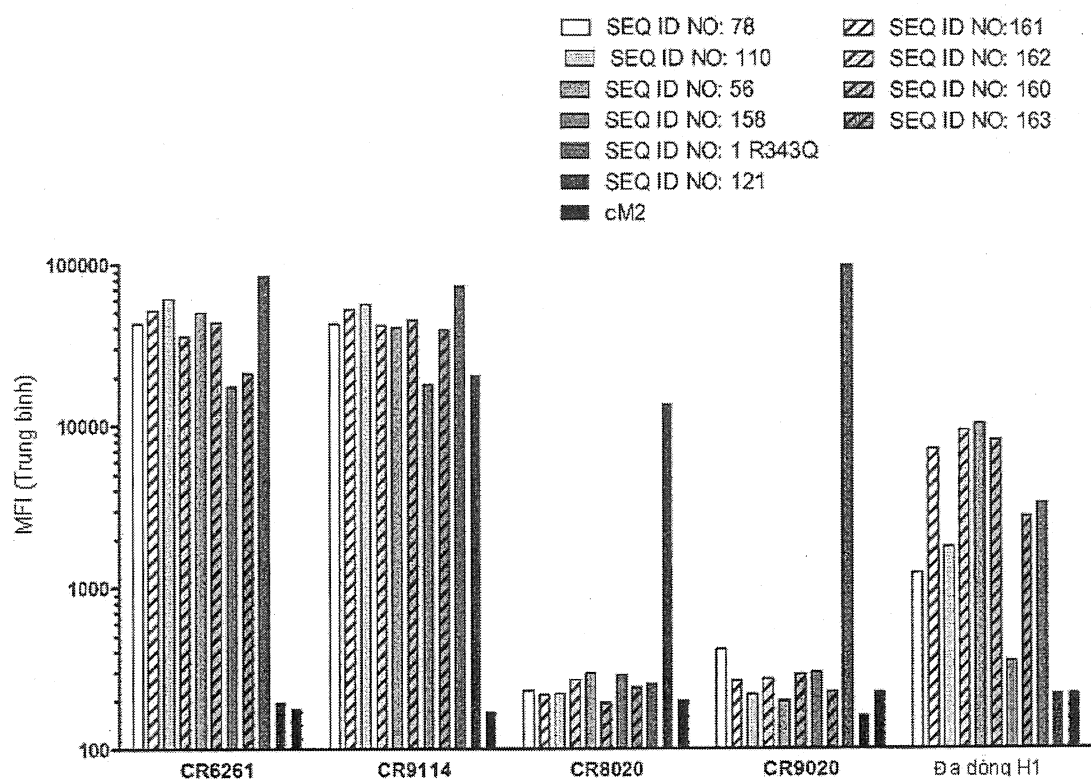
36/52

FIG. 22C



37/52

FIG. 23



38/52

FIG. 24A

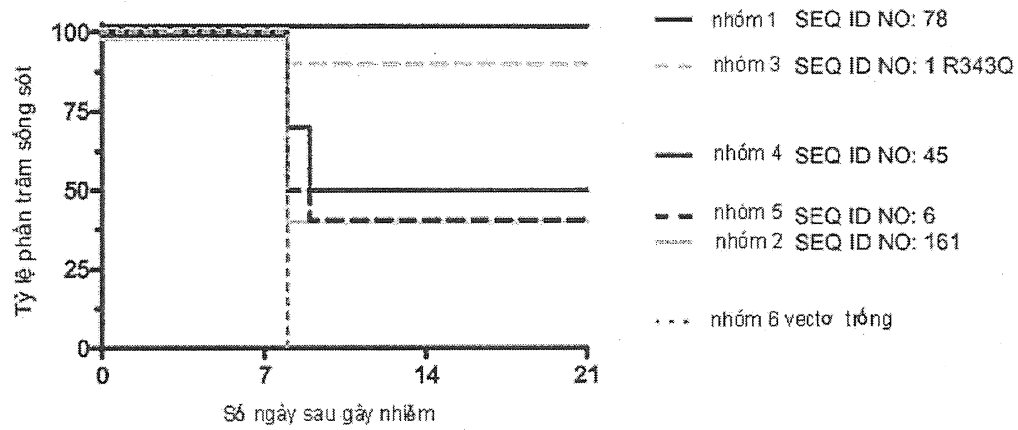
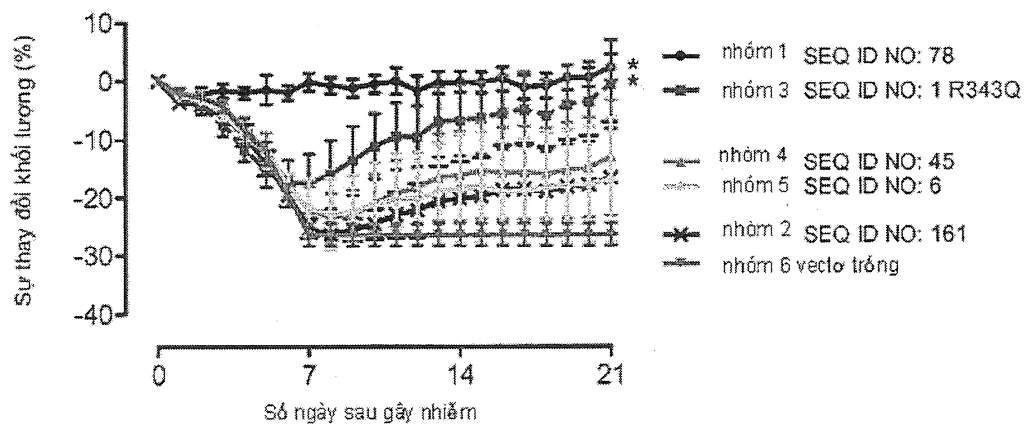


FIG. 24B



39/52

FIG. 24 C

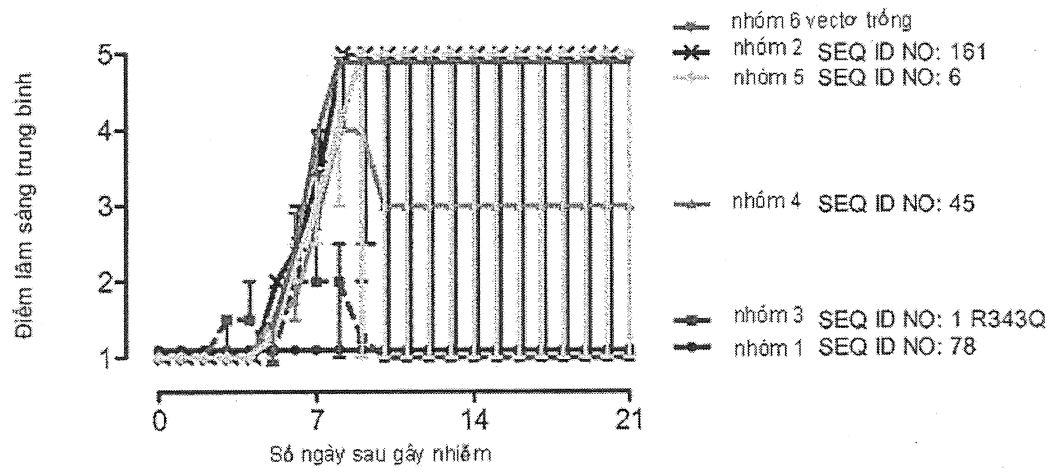
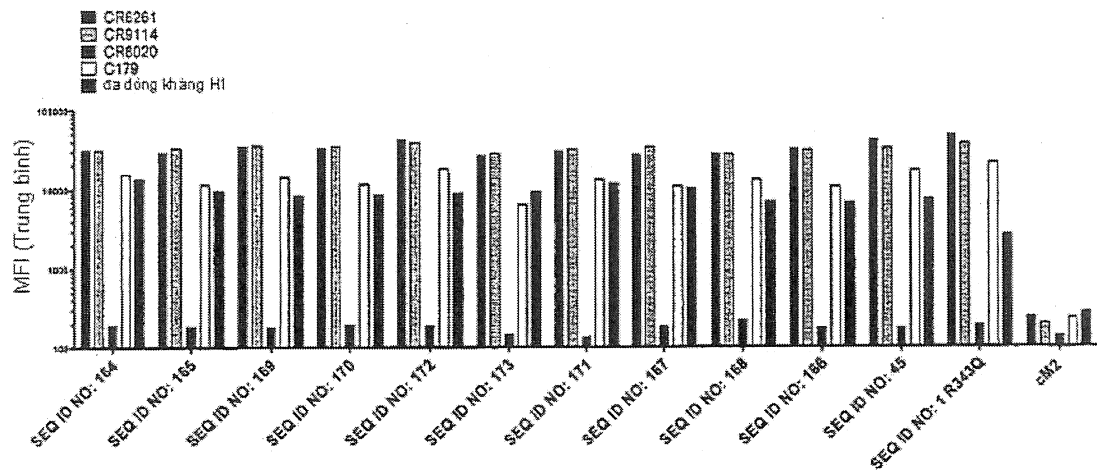


Fig. 25

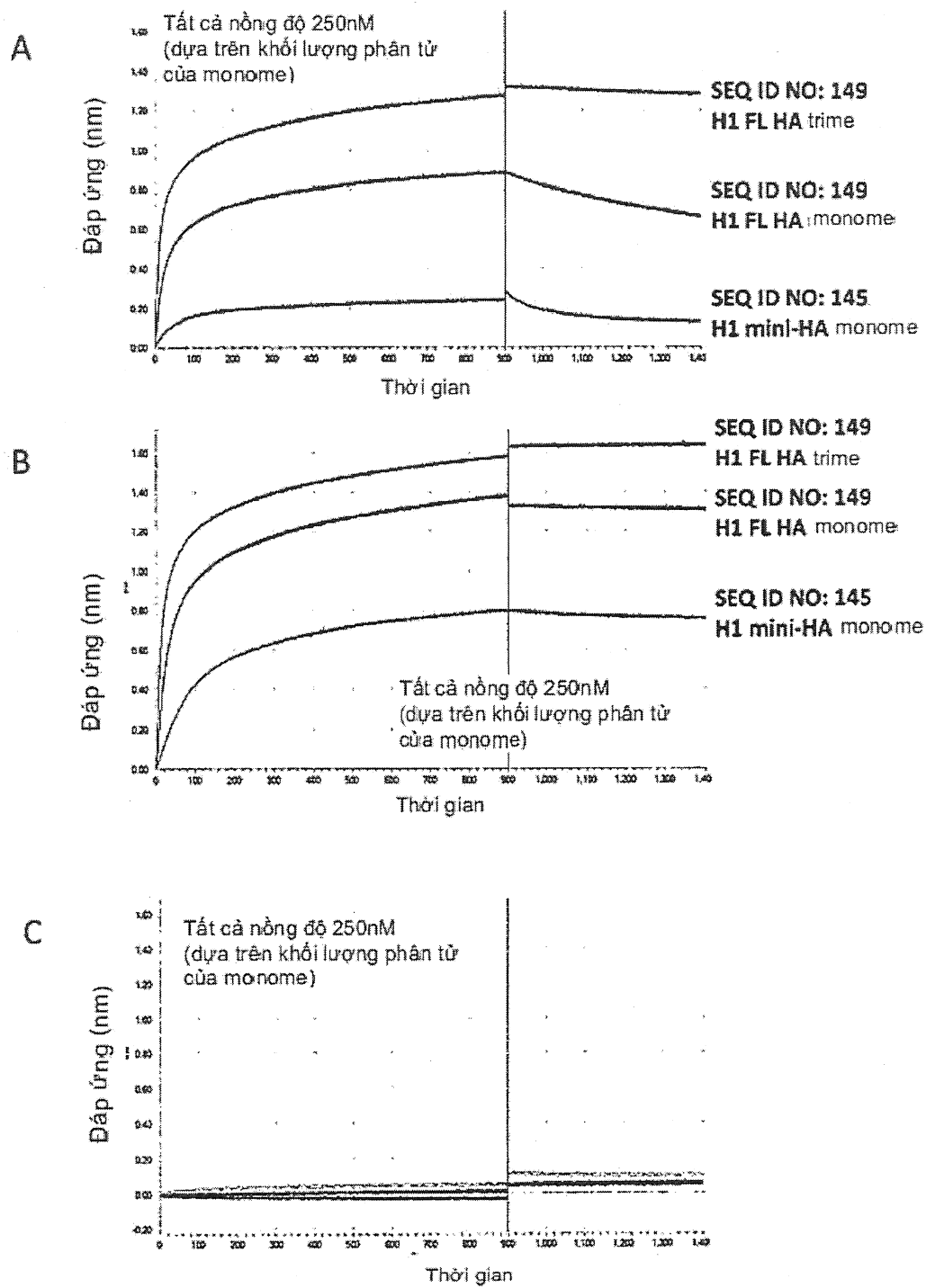
SEQ ID NO: 164: mHA HINI A_Maryland_12_1991	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190	200	210	220	230	240	250	260	270	280	290	300	310	320	330	340	350	360	370	380	390	400	410	420	430	440	450	460	470	480	490	500	510	520	530	540	550	560	570	580	590	600	610	620	630	640	650	660	670	680	690	700	710	720	730	740	750	760	770	780	790	800	810	820	830	840	850	860	870	880	890	900	910	920	930	940	950	960	970	980	990	1000	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560	1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020	2030	2040	2050	2060	2070	2080	2090	2100	2110	2120	2130	2140	2150	2160	2170	2180	2190	2200	2210	2220	2230	2240	2250	2260	2270	2280	2290	2300	2310	2320	2330	2340	2350	2360	2370	2380	2390	2400	2410	2420	2430	2440	2450	2460	2470	2480	2490	2500	2510	2520	2530	2540	2550	2560	2570	2580	2590	2600	2610	2620	2630	2640	2650	2660	2670	2680	2690	2700	2710	2720	2730	2740	2750	2760	2770	2780	2790	2800	2810	2820	2830	2840	2850	2860	2870	2880	2890	2900	2910	2920	2930	2940	2950	2960	2970	2980	2990	3000	3010	3020	3030	3040	3050	3060	3070	3080	3090	3100	3110	3120	3130	3140	3150	3160	3170	3180	3190	3200	3210	3220	3230	3240	3250	3260	3270	3280	3290	3300	3310	3320	3330	3340	3350	3360	3370	3380	3390	3400	3410	3420	3430	3440	3450	3460	3470	3480	3490	3500	3510	3520	3530	3540	3550	3560	3570	3580	3590	3600	3610	3620	3630	3640	3650	3660	3670	3680	3690	3700	3710	3720	3730	3740	3750	3760	3770	3780	3790	3800	3810	3820	3830	3840	3850	3860	3870	3880	3890	3900	3910	3920	3930	3940	3950	3960	3970	3980	3990	4000	4010	4020	4030	4040	4050	4060	4070	4080	4090	4100	4110	4120	4130	4140	4150	4160	4170	4180	4190	4200	4210	4220	4230	4240	4250	4260	4270	4280	4290	4300	4310	4320	4330	4340	4350	4360	4370	4380	4390	4400	4410	4420	4430	4440	4450	4460	4470	4480	4490	4500	4510	4520	4530	4540	4550	4560	4570	4580	4590	4600	4610	4620	4630	4640	4650	4660	4670	4680	4690	4700	4710	4720	4730	4740	4750	4760	4770	4780	4790	4800	4810	4820	4830	4840	4850	4860	4870	4880	4890	4900	4910	4920	4930	4940	4950	4960	4970	4980	4990	5000	5010	5020	5030	5040	5050	5060	5070	5080	5090	5100	5110	5120	5130	5140	5150	5160	5170	5180	5190	5200	5210	5220	5230	5240	5250	5260	5270	5280	5290	5300	5310	5320	5330	5340	5350	5360	5370	5380	5390	5400	5410	5420	5430	5440	5450	5460	5470	5480	5490	5500	5510	5520	5530	5540	5550	5560	5570	5580	5590	5600	5610	5620	5630	5640	5650	5660	5670	5680	5690	5700	5710	5720	5730	5740	5750	5760	5770	5780	5790	5800	5810	5820	5830	5840	5850	5860	5870	5880	5890	5900	5910	5920	5930	5940	5950	5960	5970	5980	5990	6000	6010	6020	6030	6040	6050	6060	6070	6080	6090	6100	6110	6120	6130	6140	6150	6160	6170	6180	6190	6200	6210	6220	6230	6240	6250	6260	6270	6280	6290	6300	6310	6320	6330	6340	6350	6360	6370	6380	6390	6400	6410	6420	6430	6440	6450	6460	6470	6480	6490	6500	6510	6520	6530	6540	6550	6560	6570	6580	6590	6600	6610	6620	6630	6640	6650	6660	6670	6680	6690	6700	6710	6720	6730	6740	6750	6760	6770	6780	6790	6800	6810	6820	6830	6840	6850	6860	6870	6880	6890	6900	6910	6920	6930	6940	6950	6960	6970	6980	6990	7000	7010	7020	7030	7040	7050	7060	7070	7080	7090	7100	7110	7120	7130	7140	7150	7160	7170	7180	7190	7200	7210	7220	7230	7240	7250	7260	7270	7280	7290	7300	7310	7320	7330	7340	7350	7360	7370	7380	7390	7400	7410	7420	7430	7440	7450	7460	7470	7480	7490	7500	7510	7520	7530	7540	7550	7560	7570	7580	7590	7600	7610	7620	7630	7640	7650	7660	7670	7680	7690	7700	7710	7720	7730	7740	7750	7760	7770	7780	7790	7800	7810	7820	7830	7840	7850	7860	7870	7880	7890	7900	7910	7920	7930	7940	7950	7960	7970	7980	7990	8000	8010	8020	8030	8040	8050	8060	8070	8080	8090	8100	8110	8120	8130	8140	8150	8160	8170	8180	8190	8200	8210	8220	8230	8240	8250	8260	8270	8280	8290	8300	8310	8320	8330	8340	8350	8360	8370	8380	8390	8400	8410	8420	8430	8440	8450	8460	8470	8480	8490	8500	8510	8520	8530	8540	8550	8560	8570	8580	8590	8600	8610	8620	8630	8640	8650	8660	8670	8680	8690	8700	8710	8720	8730	8740	8750	8760	8770	8780	8790	8800	8810	8820	8830	8840	8850	8860	8870	8880	8890	8900	8910	8920	8930	8940	8950	8960	8970	8980	8990	9000	9010	9020	9030	9040	9050	9060	9070	9080	9090	9100	9110	9120	9130	9140	9150	9160	9170	9180	9190	9200	9210	9220	9230	9240	9250	9260	9270	9280	9290	9300	9310	9320	9330	9340	9350	9360	9370	9380	9390	9400	9410	9420	9430	9440	9450	9460	9470	9480	9490	9500	9510	9520	9530	9540	9550	9560	9570	9580	9590	9600	9610	9620	9630	9640	9650	9660	9670	9680	9690	9700	9710	9720	9730	9740	9750	9760	9770	9780	9790	9800	9810	9820	9830	9840	9850	9860	9870	9880	9890	9900	9910	9920	9930	9940	9950	9960	9970	9980	9990	10000	10010	10020	10030	10040	10050	10060	10070	10080	10090	10100	10110	10120	10130	10140	10150	10160	10170	10180	10190	10200	10210	10220	10230	10240	10250	10260	10270	10280	10290	10300	10310	10320	10330	10340	10350	10360	10370	10380	10390	10400	10410	10420	10430	10440	10450	10460	10470	10480	10490	10500	10510	10520	10530	10540	10550	10560	10570	10580	10590	10600	10610	10620	10630	10640	10650	10660	10670	10680	10690	10700	10710	10720	10730	10740	10750	10760	10770	10780	10790	10800	10810	10820	10830	10840	10850	10860	10870	10880	10890	10900	10910	10920	10930	10940	10950	10960	10970	10980	10990	11000	11010	11020	11030	11040	11050	11060	11070	11080	11090	11100	11110	11120	11130	11140	11150	11160	11170	11180	11190	11200	11210	11220	11230	11240	11250	11260	11270	11280	11290	11300	11310	11320	11330	11340	11350	11360	11370	11380	11390	11400	11410	11420	11430	11440	11450	11460	11470	11480	11490	11500	11510	11520	11530	11540	11550	11560	11570	11580	11590	11600	11610	11620	11630	11640	11650	11660	11670	11680	11690	11700	11710	11720	11730	11740	11750	11760	11770	11780	11790	11800	11810	11820	11830	11840	11850	11860	11870	11880	11890	11900	11910	11920	11930	11940	11950	11960	11970	11980	11990	12000	12010	12020	12030	12040	12050	12060	12070	12080	12090	12100	12110	12120	12130	12140	12150	12160	12170	12180	12190	12200	12210	12220	12230	12240	12250	12260	12270	12280	12290	12300	12310	12320	12330	12340	12350	12360	12370	12380	12390	12400	12410	12420	12430	12440	12450	12460	12470	12480	12490	12500	12510	12520	12530	12540	12550	12560	12570	12580	12590	12600	12610	12620	12630	12640	12650	12660	12670	12680	12690	12700	12710	12720	12730	12740	12750	12760	12770	12780	12790	12800	12810	12820	12830	12840	12850	12860	12870	12880	12890	12900	12910	12920	12930	12940	12950	12960	12970	12980	12990	13000	13010	13020	13030	13040	13050	13060	13070	13080	13090	13100	13110	13120	13130	13140	13150	13160	13170	13180	13190	13200	13210	13220	13230	13240	13250	13260	13270	13280	13290	13300	13310	13320
---	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

FIG. 26



42/52

FIG. 27



43/52

FIG. 28

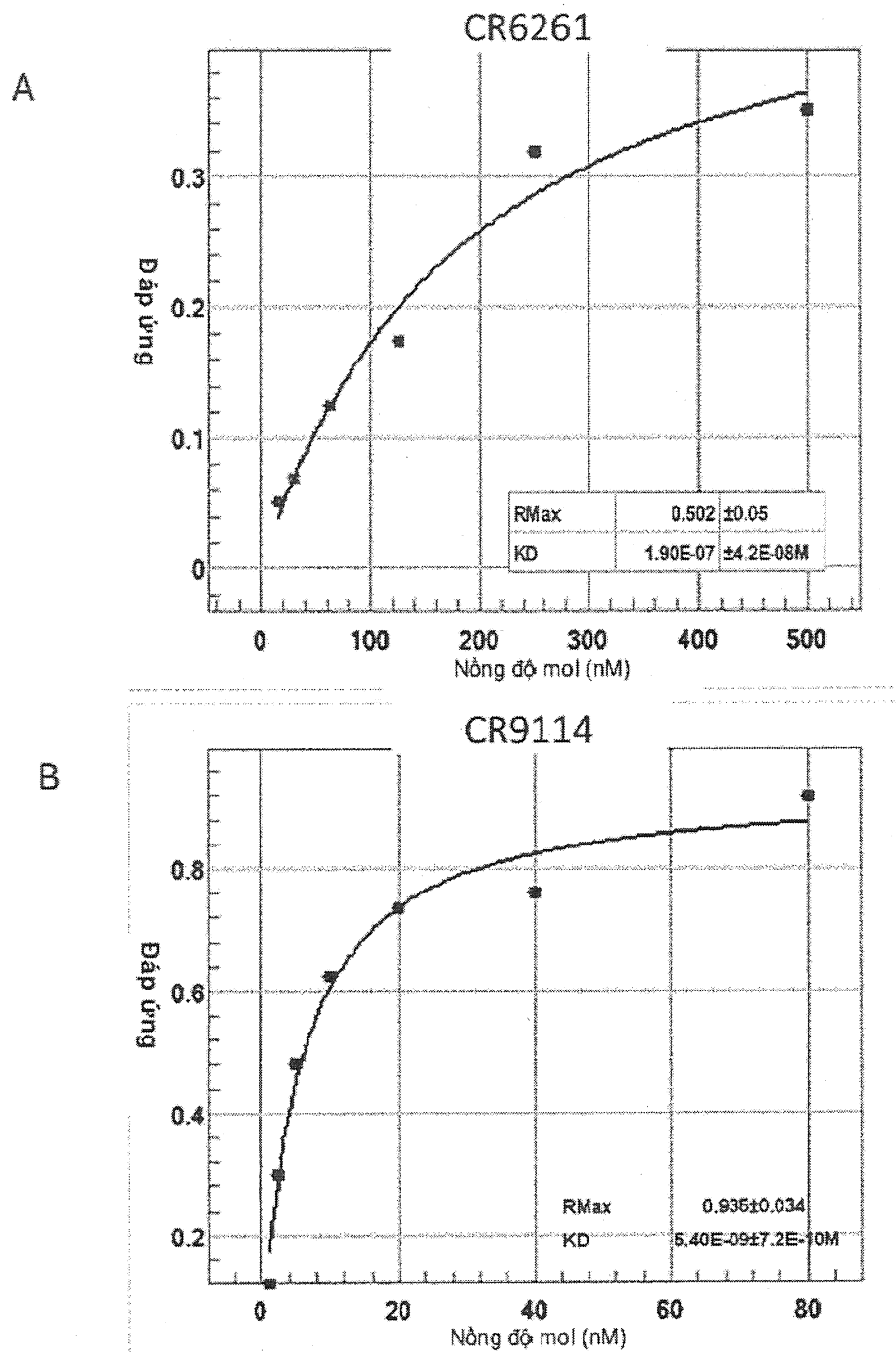
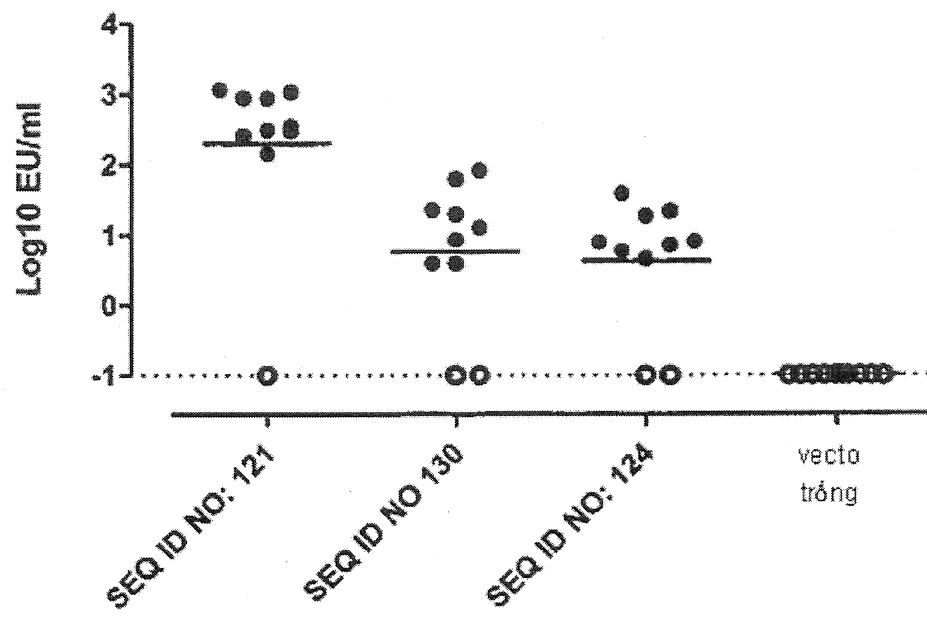


FIG. 29



vỏ ngoài : rHA A/HK/1/68

45/52

FIG. 30A

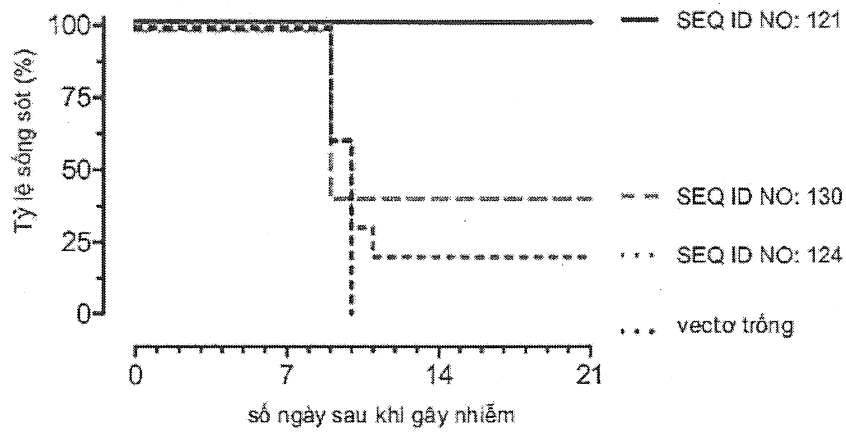
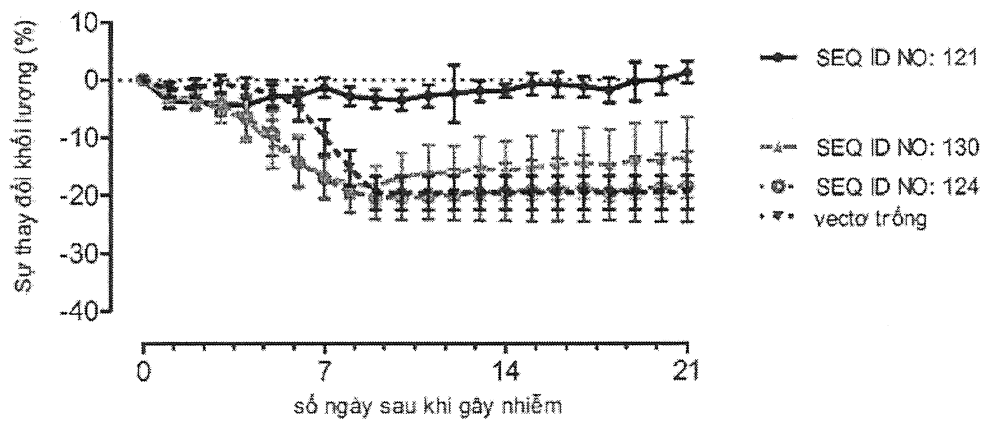


FIG. 30B



46/52

FIG. 30C

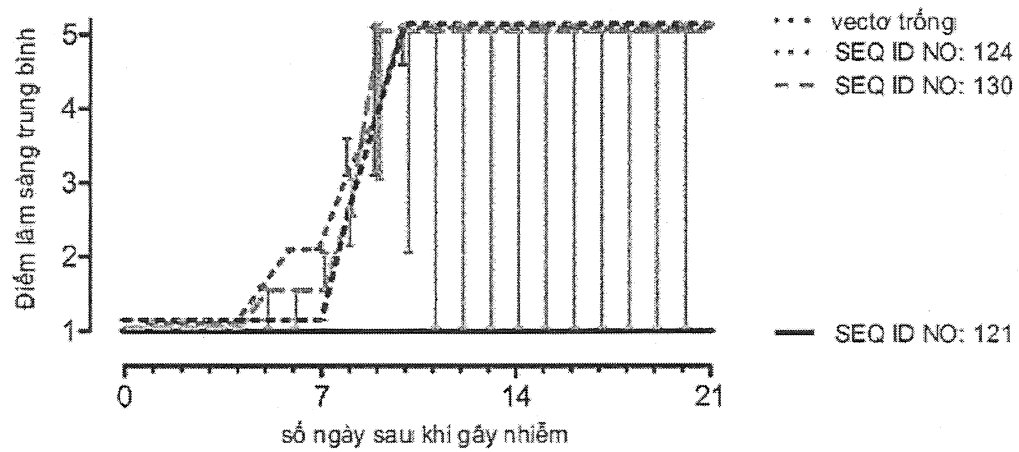
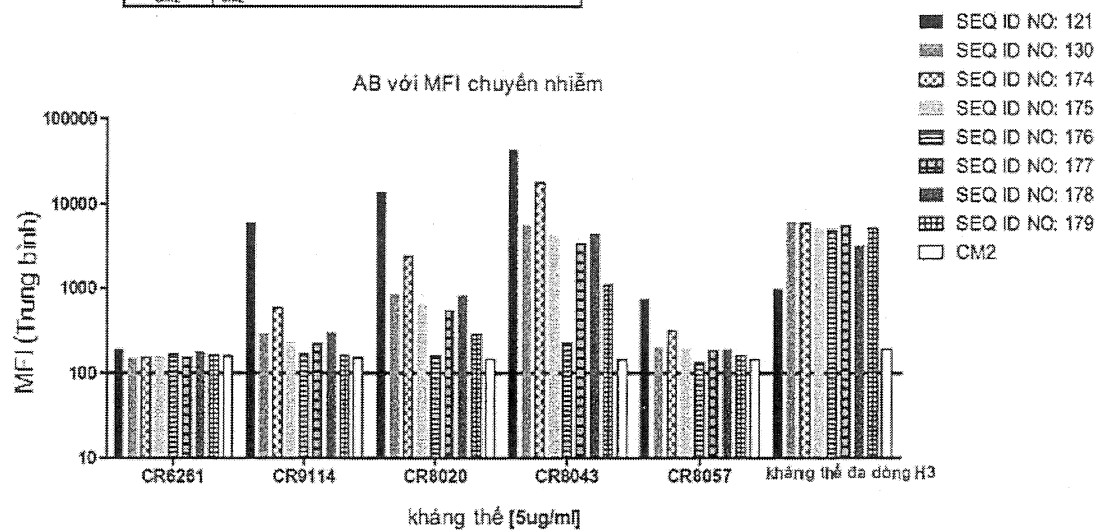


FIG. 31

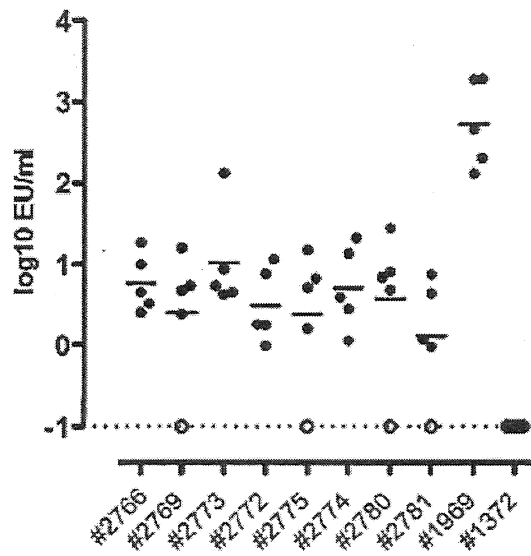
SEQ ID NO	Tên
121	pcDNA2004Neo_H3-Hongkong 88 R345Q
130	pcDNA2004Neo_H3min2a- Cầu nối +c9+10+11+12-GCN4
174	H3 HK min2a- Cầu nối +c9+10+11+12+GCN4T-CGT-1
175	HK68 H3min2a- Cầu nối +c9+10+12+18+GCN4T
176	HK68 H3min2a- Cầu nối +c9+10+12+16+CG7-GCN4T
177	HK68 H3min2a- Cầu nối +c9+10+12+19+GCN4T
178	HK68 H3min2a- Cầu nối +c9+10+12+17+CG7-GCN4T
179	H3 HK68 min2a- Cầu nối 2+c9+10+12+GCN4T
CM2	CM2



48/52

FIG. 32A

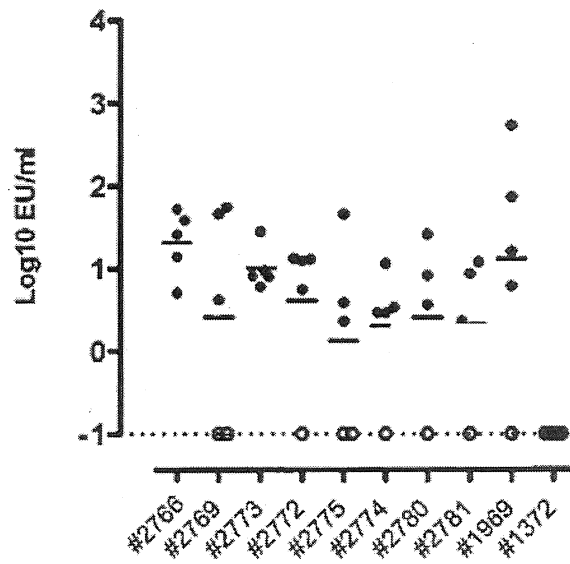
#2766 SEQ ID NO: 105: H3-mini2
 #2769 SEQ ID NO: 108: H3-mini2-cl9+10+11
 #2773 SEQ ID NO: 112: H3-mini2-cl9+10+12
 #2772 SEQ ID NO: 111: H3-mini2-cl9+10+11+12
 #2775 SEQ ID NO: 114: H3-mini2-cl9+10+11+12-tri
 #2774 SEQ ID NO: 113: H3-mini2-cl9+10+11+12-GCN4
 #2780 SEQ ID NO: 119: H3-mini3-cl9+10+11+12+14
 #2781 SEQ ID NO: 120: H3-mini4-cl9+10+11+12+14
 #1969 SEQ ID NO: 89 H3 chiều dài đủ A/Wisconsin/67/2005
 #1372 cM2



49/52

FIG. 32B

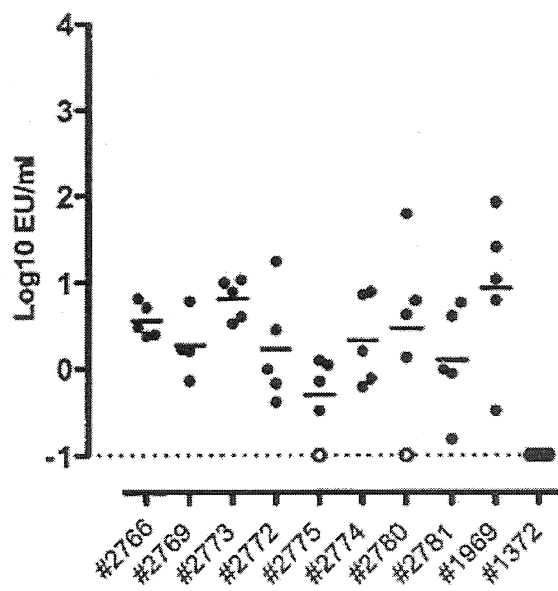
#2766 SEQ ID NO: 105: H3-mini2
 #2769 SEQ ID NO: 108: H3-mini2-cl9+10+11
 #2773 SEQ ID NO: 112: H3-mini2-cl9+10+12
 #2772 SEQ ID NO: 111: H3-mini2-cl9+10+11+12
 #2775 SEQ ID NO: 114: H3-mini2-cl9+10+11+12-tri
 #2774 SEQ ID NO: 113: H3-mini2-cl9+10+11+12-GCN4
 #2780 SEQ ID NO: 119: H3-mini3-cl9+10+11+12+14
 #2781 SEQ ID NO: 120: H3-mini4-cl9+10+11+12+14
 #1969 SEQ ID NO: 89 H3 chiều dài đủ A/Wisconsin/67/2005
 #1372 cM2



50/52

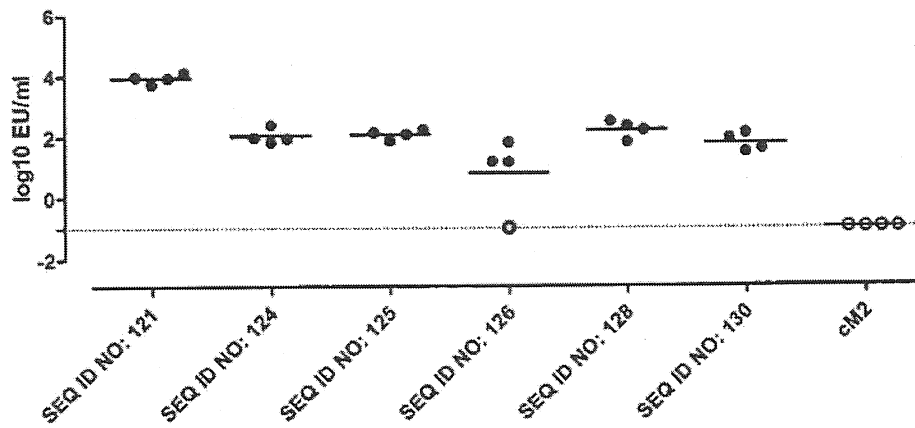
FIG. 32C

#2766 SEQ ID NO: 105: H3-mini2
 #2769 SEQ ID NO: 108: H3-mini2-cl9+10+11
 #2773 SEQ ID NO: 112: H3-mini2-cl9+10+12
 #2772 SEQ ID NO: 111: H3-mini2-cl9+10+11+12
 #2775 SEQ ID NO: 114: H3-mini2-cl9+10+11+12-tri
 #2774 SEQ ID NO: 113: H3-mini2-cl9+10+11+12-GCN4
 #2780 SEQ ID NO: 119: H3-mini3-cl9+10+11+12+14
 #2781 SEQ ID NO: 120: H3-mini4-cl9+10+11+12+14
 #1969 SEQ ID NO: 89 H3 chiều dài đủ A/Wisconsin/67/2005
 #1372 cM2



51/52

FIG. 33



vỏ ngoài: rHA A/HongKong/1/68 (có nguồn gốc từ tế bào người)

52/52

FIG. 34

