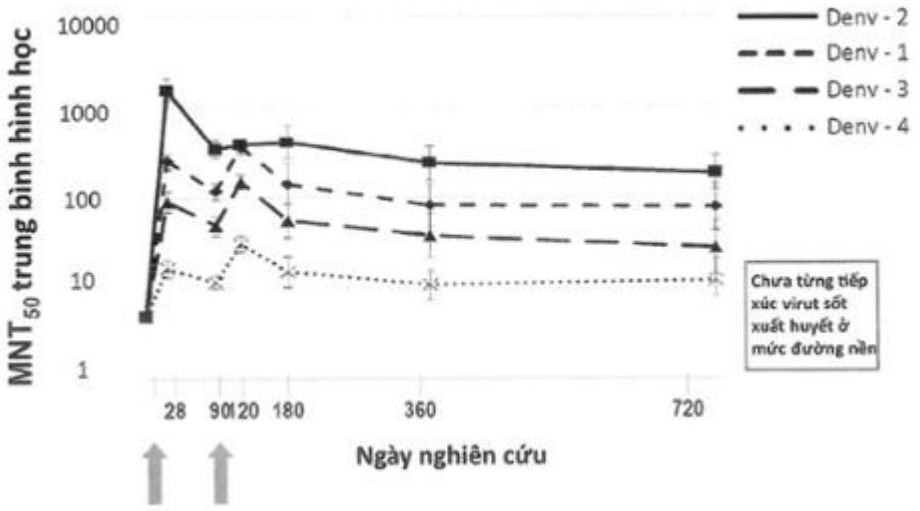
(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) **Cộng hòa xã hội chủ nghĩa Việt Nam (VN)**
CỤC SỞ HỮU TRÍ TUỆ
(11) **1-0038096**
(51)[8] **A61K 39/12**; A61P 31/14
(13) **B**

(21) 1-2018-04958
(22) 13/04/2017
(86) PCT/IB2017/052160 13/04/2017
(87) WO2017/179017 19/10/2017
(30) 62/322,167 13/04/2016 US
(45) 25/01/2024 430
(43) 27/05/2019 374A
(73) TAKEDA VACCINES, INC. (US)
75 Sidney Street, Cambridge, MA 02139, United States of America
(72) WALLACE, Derek (GB); BOSLEGO, John (US).
(74) Công ty Luật TNHH T&G (TGVN)

(54) DƯỢC PHẨM CHỨA CHẾ PHẨM VACXIN TỨ GIÁ ĐỂ ĐIỀU TRỊ CHO TRẺ EM HOẶC THANH NIÊN CHỐNG LẠI SỰ NHIỄM VIRUT SỐT XUẤT HUYẾT

(57) Sáng chế đề cập đến dược phẩm chứa ít nhất một virut sốt xuất huyết sống đã được giảm độc lực và thể khảm virut sốt xuất huyết-virut sốt xuất huyết để điều trị cho trẻ em hoặc thanh niên chống lại sự nhiễm virut sốt xuất huyết, trong đó dược phẩm này gây ra đáp ứng miễn dịch đối với virut sốt xuất huyết ở trẻ em hoặc thanh niên. Sáng chế cũng đề cập đến dược phẩm chứa ít nhất một virut sốt xuất huyết sống đã được giảm độc lực và thể khảm virut sốt xuất huyết-virut sốt xuất huyết để điều trị cho trẻ em từ 1,5 đến 11 tuổi chống lại sự nhiễm virut sốt xuất huyết, trong đó dược phẩm này gây ra đáp ứng miễn dịch đối với virut sốt xuất huyết ở trẻ em.



Lĩnh vực kỹ thuật được đề cập

Các phương án được bộc lộ ở đây liên quan đến chế phẩm, phương pháp, và các ứng dụng để gây ra đáp ứng miễn dịch đối với cả bốn typ huyết thanh virut sốt xuất huyết ở trẻ em hoặc thanh niên từ khoảng 1 tuổi đến khoảng 20 tuổi. Một số phương án liên quan đến chế phẩm mà có thể bao gồm, nhưng không giới hạn ở, các cấu trúc virut flavivirut khảm và không khảm mà, hoặc ở dạng riêng rẽ hoặc kết hợp, có thể được sử dụng trong chế phẩm vacxin để gây đáp ứng miễn dịch chống lại cả bốn typ huyết thanh virut sốt xuất huyết. Theo các phương án nhất định, chế phẩm có thể bao gồm các cấu trúc gồm nhiều hơn một typ huyết thanh của virut sốt xuất huyết, như virut sốt xuất huyết-1 (DEN-1), virut sốt xuất huyết-2 (DEN-2), virut sốt xuất huyết-3 (DEN-3) và/hoặc virut sốt xuất huyết-4 (DEN-4), ở các nồng độ hoặc tỷ lệ khác nhau để cải thiện sự bảo vệ khỏi lây nhiễm sốt xuất huyết ở trẻ em và thanh niên. Các phương án khác liên quan đến phương pháp sử dụng chế phẩm vacxin mà có thể bao gồm các cấu trúc virut sốt xuất huyết thể khảm và virut sốt xuất huyết sống giảm độc lực bằng cách sử dụng phác đồ sử dụng đơn, kép hoặc phác đồ khác.

Tình trạng kỹ thuật của sáng chế

Sốt xuất huyết là một bệnh được lây bởi muỗi gây ra do sự nhiễm virut sốt xuất huyết. Nhiễm virut sốt xuất huyết có thể dẫn đến các triệu chứng suy nhược và đau đớn, bao gồm sốt cao đột ngột, đau đầu, đau khớp và cơ, buồn nôn, nôn và phát ban ở da. Cho đến nay, bốn typ huyết thanh của virut sốt xuất huyết đã được nhận diện: virut sốt xuất huyết-1 (DEN-1), virut sốt xuất huyết-2 (DEN-2), hoặc virut sốt xuất huyết-3 (DEN-3) kết hợp với virut sốt xuất huyết-4 (DEN-4). Các kiểu phụ khác có thể được phát hiện trong tương lai (*chẳng hạn*, DEN-5). Các typ huyết thanh virut sốt xuất huyết 1-4 cũng có thể gây ra sốt xuất huyết do virut sốt xuất huyết (dengue hemorrhagic fever - DHF) và hội chứng sốc do virut sốt xuất huyết (dengue shock syndrome - DSS). Trong hầu hết các trường hợp nghiêm trọng, DHF và DSS

có thể gây chết người. Các virut sốt xuất huyết gây ra 50-100 triệu trường hợp sốt xuất huyết suy nhược, 500.000 trường hợp DHF/DSS, và nhiều hơn 20.000 trường hợp tử vong mỗi năm, phần lớn trong số chúng là trẻ em. Cho đến nay, không có vacxin hữu hiệu để bảo vệ chống sốt xuất huyết và không có phương pháp điều trị bằng thuốc nào đối với bệnh này. Các nỗ lực nhằm kiểm soát muỗi là không có hiệu quả trong ngăn ngừa sự bùng phát virut sốt xuất huyết ở các khu vực địa phương hoặc trong ngăn ngừa sự lan rộng về mặt địa lý của bệnh này. Ước tính rằng 3,5 tỷ người bị đe dọa lây nhiễm virut sốt xuất huyết. Ngoài ra, virut sốt xuất huyết là nguyên nhân hàng đầu gây ra sốt đối với khách du lịch đến các khu vực địa phương, như Châu Á, trung tâm Nam mỹ, và vùng Ca-ri-bê. Thêm vào đó, DHF/DSS là nguyên nhân hàng đầu dẫn đến ốm nghiêm trọng và tử vong ở trẻ em tại một số nước Châu Á và Châu Mỹ La tinh.

Cả bốn typ huyết thanh virut sốt xuất huyết đều mang tính địa phương trên toàn bộ vùng nhiệt đới trên thế giới và tạo ra mối đe dọa nghiêm trọng nhất do virut được mang bởi muỗi gây ra cho loài người ở các vùng nhiệt đới, trên toàn thế giới. Virut sốt xuất huyết được truyền cho người chủ yếu bởi muỗi *Aedes aegypti*. Nhiễm một typ huyết thanh virut sốt xuất huyết dẫn đến sự bảo vệ trong thời gian dài khỏi sự nhiễm lại của typ huyết thanh đó, nhưng không bảo vệ được sự nhiễm lần thứ hai bởi một trong số ba typ huyết thanh virut sốt xuất huyết khác. Trên thực tế, lây nhiễm trước với một typ huyết thanh virut sốt xuất huyết dẫn đến nguy cơ gia tăng bệnh nghiêm trọng (DHF/DSS) do sự nhiễm lần thứ hai với một typ huyết thanh khác. Việc phát triển một vacxin hữu hiệu, đặc biệt là ở trẻ em và thanh niên, là một cách tiếp cận quan trọng để ngăn ngừa và kiểm soát bệnh này trên toàn cầu.

Bản chất kỹ thuật của sáng chế

Các phương án được bộc lộ ở đây liên quan đến chế phẩm, phương pháp và các ứng dụng của cấu trúc virut sốt xuất huyết thể khảm. Theo một số phương án, chế phẩm có thể bao gồm cấu trúc virut sốt xuất huyết thể khảm có bộ khung virut sốt xuất huyết đã được giảm độc lực với các gen cấu trúc từ ít nhất một typ huyết thanh virut sốt xuất huyết khác một mình hoặc kết hợp với virut sốt xuất huyết sống giảm độc lực. Các phương án khác liên quan đến ít nhất một virut sống đã được giảm

độc lực kết hợp với một hoặc nhiều virut sốt xuất huyết khảm chứa các phần tử cấu tạo chống lại ít nhất một typ huyết thanh virut sốt xuất huyết khác. Các phương án khác có thể bao gồm chế phẩm chứa virut sốt xuất huyết khảm có bộ khung DEN-2 cải biến (*chẳng hạn*, PDK-53 là bộ khung khởi đầu ở P1 (cấy chuyển-1) và PDK-53 cải biến bằng cách sử dụng tính biến thiên qua cấy chuyển và chọn lọc như đã chỉ ra đối với P2, P3...P8..P10 v.v) và một hoặc nhiều thành phần cấu tạo của DEN-1, DEN-2, DEN-3 hoặc DEN-4. Theo một số phương án, chế phẩm sinh miễn dịch có thể được tạo ra ở đó khi được đưa vào đối tượng, chế phẩm này tạo ra đáp ứng miễn dịch đối với một hoặc nhiều virut sốt xuất huyết ở đối tượng. Vì vậy, các cấu trúc được dự tính ở đây có thể được tạo ra và được cấy chuyển và mỗi lần cấy chuyển tạo ra một virut sốt xuất huyết giảm độc lực được dự tính để dùng trong chế phẩm vacxin dược dụng.

Các phương án được bộc lộ ở đây liên quan đến chế phẩm, phương pháp, và các ứng dụng để gây ra đáp ứng miễn dịch đối với cả bốn typ huyết thanh virut sốt xuất huyết ở trẻ em hoặc thanh niên từ khoảng 1 tuổi đến khoảng 45 tuổi. Các chế phẩm hiện nay là không hữu hiệu đối với thanh niên và luôn có nhu cầu đối với vacxin virut sốt xuất huyết dùng để chống lại nhiễm virut sốt xuất huyết đối với tất cả các nhóm tuổi. Theo các phương án này, chế phẩm và phương pháp được bộc lộ ở đây là để dùng cho trẻ em tuổi từ khoảng 1 đến 20 tuổi, từ khoảng 1 đến 15 tuổi, từ khoảng 1 đến 11 tuổi hoặc từ khoảng 2 đến khoảng 9 tuổi hoặc từ khoảng 1 đến khoảng 5 tuổi.

Theo các phương án nhất định, cấu trúc virut sốt xuất huyết thể khảm của typ huyết thanh virut sốt xuất huyết để dùng trong vacxin có thể bao gồm virut sống đã được giảm độc lực đã được cấy chuyển 8 lần (P8) hoặc virut thể khảm có trình tự axit nucleic được nhận diện bởi SEQ ID NOS: 1, 3, 5 và 7 hoặc trình tự polypeptit được nhận diện bởi SEQ ID NOS: 2, 4, 6 và 8. Ở đây dự tính rằng dạng cấy chuyển bất kỳ trong số các dạng cấy chuyển đối với bất kỳ trong số các virut sống đã được giảm độc lực được mô tả ở đây có thể được sử dụng trong chế phẩm sinh miễn dịch để gây ra đáp ứng miễn dịch đối với typ huyết thanh virut sốt xuất huyết điển hình (*chẳng hạn*, các typ huyết thanh 1-4). Theo các phương án này, chế phẩm sinh miễn

dịch bao gồm virut sống đã được giảm độc lực đã được phân lập P-8 có thể được dùng cho đối tượng để gây ra đáp ứng sinh miễn dịch đối với một hoặc nhiều typ huyết thanh virut sốt xuất huyết tùy thuộc vào cấu trúc được lựa chọn. Ngoài ra, virut sống đã được giảm độc lực có thể được kết hợp với một hoặc nhiều virut thể khảm này. Điều này được dự tính đối với mỗi trong số các virut sống đã được giảm độc lực được phân lập/được tạo ra trong lần cấy chuyển tế bào tiếp theo (*e.g.*, African Green Monkey Vero cell production, hereinafter: Vero cells or other appropriate cell line). Bản mô tả này dự định là dòng tế bào bất kỳ (*chẳng hạn*, được phê chuẩn GMP) có khả năng sản xuất virut sốt xuất huyết được sử dụng để cấy chuyển bất kỳ trong số các cấu trúc virut ở quy mô sản xuất hoặc khi thích hợp bản mô tả này dự tính việc sử dụng tiếp theo vacxin hoặc chế phẩm sinh miễn dịch chống lại virut sốt xuất huyết.

Theo các phương án khác, các chế phẩm được dự tính ở đây có thể được kết hợp với các chế phẩm sinh miễn dịch khác chống lại các Flavivirut khác như virut Zika, virut West Nile, virut gây viêm não Nhật Bản, virut gây viêm não St. Louis, virut gây sốt vàng da hoặc cấu trúc khảm flavivirut bất kỳ khác và/hoặc virut sống đã được giảm độc lực. Theo các phương án nhất định, chế phẩm đơn lẻ có thể được sử dụng để chống lại nhiều flavivirut. Theo một số phương án, chế phẩm được bộc lộ ở đây có thể có khả năng gây đáp ứng miễn dịch chống lại virut Zika ngay cả khi chế phẩm vacxin bị giới hạn ở các cấu trúc virut sốt xuất huyết và virut sốt xuất huyết sống đã được giảm độc lực.

Theo các phương án nhất định, chế phẩm sinh miễn dịch theo sáng chế có thể bao gồm virut sốt xuất huyết khảm chống lại một hoặc nhiều trong số DEN-1, DEN-2, DEN-3 và/hoặc DEN-4, một mình hoặc kết hợp với chế phẩm virut sốt xuất huyết sống đã được giảm độc lực.

Theo một phương án khác, cấu trúc có thể bao gồm cấu trúc mang các đột biến thích nghi ở vùng cấu trúc hoặc không cấu trúc của virut này mà làm tăng sự sinh trưởng hoặc sự sản sinh mà không làm ảnh hưởng đến sự giảm độc lực hoặc độ an toàn của virut khi được đưa vào đối tượng. Theo các phương án nhất định, cấu trúc virut sốt xuất huyết thể khảm được dự tính bất kỳ có thể bao gồm virut DEN-2 sống giảm độc lực mang các đột biến đặc hiệu được dùng làm bộ khung mà ở đó

virut DEN-2 PDK sống giảm độc lực có thể còn bao gồm protein cấu trúc gồm một hoặc nhiều protein cấu trúc prM (tiền màng) và E (lớp vỏ) của các typ huyết thanh virut sốt xuất huyết khác. Ngoài ra, bộ khung DEN-2 có thể bao gồm các đột biến bổ sung hoặc đột biến ngược để tăng cường đáp ứng miễn dịch đối với một chế phẩm định trước ở đối tượng ngay khi sử dụng (*chẳng hạn*, virut sốt xuất huyết thể khảm 2/1, 2/3 hoặc 2/4).

Theo một số phương án, các gen protein cấu trúc có thể bao gồm các gen prM và E của DEN-1, DEN-2, DEN-3 hoặc DEN-4 trên bộ khung DEN-2 chứa một hoặc hai đột biến ngược để cải thiện tính sinh miễn dịch. Ví dụ, cấu trúc virut sốt xuất huyết, theo các phương án nhất định có thể bao gồm các cấu trúc được gọi là TDV-1-A, TDV-2-F, TDV-3-F, và TDV-4-F (xem phần Ví dụ) trong đó bộ khung DEN-2 có một hoặc nhiều đột biến đảo ngược thành DEN-2 kiểu dại (*chẳng hạn*, trong vùng không mã hóa (NCR) hoặc vùng không cấu trúc (NS1 v.v..) hoặc các đột biến bổ sung không được phát hiện trong P1 hoặc virut được cấy chuyển trước đó) từ virut sống đã được giảm độc lực DEN-2 được thể hiện trước đó là an toàn và hữu hiệu để gây ra đáp ứng miễn dịch. Virut sống đã được giảm độc lực DEN-2 của đơn sáng chế này là phiên bản cải tiến của virut sống đã được giảm độc lực DEN-2 được sử dụng ban đầu. Cấu trúc khảm của một số phương án được bộc lộ ở đây có thể bao gồm bộ khung DEN-2 PDK-53 được giảm độc lực cải biến, mang một hoặc nhiều protein cấu trúc của typ huyết thanh virut sốt xuất huyết thứ hai trong đó protein cấu trúc có thể bao gồm các đột biến bổ sung để tăng đáp ứng sinh miễn dịch đối với cấu trúc khảm. Theo một số phương án, các đột biến nhất định thu được bởi DEN-2 PDK-53 giảm độc lực có thể được đảo ngược thành axit amin đối chứng hoặc axit amin khác để tạo ra cấu trúc khảm khác với cấu trúc P1 mà có thể tạo ra tính sinh miễn dịch gia tăng, sự sinh trưởng gia tăng, kích thước vết tan gia tăng mà không làm ảnh hưởng đến độ an toàn và sự giảm độc lực của vacxin và có thể ảnh hưởng đến sự sinh trưởng và/hoặc sự sao chép của virut sống giảm độc lực.

Theo một phương án khác, bộ gen của DEN-2 sống giảm độc lực có thể được sử dụng để tạo ra các cấu trúc khảm của typ huyết thanh 1 virut sốt xuất huyết (DEN-1), typ huyết thanh 3 virut sốt xuất huyết (DEN-3), và typ huyết thanh 4 virut

sốt xuất huyết (DEN-4) trong đó một hoặc nhiều gen protein cấu trúc của bộ gen virut DEN-2 tương ứng có thể được thay bằng một hoặc nhiều gen protein cấu trúc của DEN-1, DEN-3 hoặc DEN-4. Theo một số phương án, protein cấu trúc có thể bao gồm protein C, prM hoặc E của virut sốt xuất huyết thứ ba. Theo các phương án nhất định, các gen protein cấu trúc có thể bao gồm các gen prM và E của DEN-1, DEN-3 hoặc DEN-4. Theo các phương án này, các virut thể khảm có thể biểu hiện các kháng nguyên bề mặt của DEN-1, DEN-3 hoặc DEN-4 trong khi vẫn duy trì kiểu hình giảm độc lực của DEN-2 bố mẹ đã được giảm độc lực. Ngoài ra, các cấu trúc khảm được bộc lộ ở đây để dùng cho các vacxin ở trẻ em và thanh niên có thể bao gồm cấu trúc khảm của DEN-4, DEN-2, DEN-1, và DEN-3 biểu hiện các kháng nguyên bề mặt của DEN-1, DEN-3 và DEN-4 bằng cách sử dụng virut DEN-2 giảm độc lực làm bộ khung.

Theo các phương án nhất định, chế phẩm theo sáng chế có thể bao gồm chế phẩm mà có thể bao gồm cấu trúc virut sốt xuất huyết khảm duy nhất được bộc lộ trong bản mô tả này và chất mang hoặc tá dược dược dụng. Theo một phương án khác, chế phẩm được bộc lộ ở đây có thể bao gồm hai hoặc nhiều hơn hai, hoặc ba hoặc nhiều hơn ba cấu trúc virut sốt xuất huyết thể khảm, và chất mang hoặc tá dược dược dụng. Theo các phương án này, một hoặc nhiều cấu trúc khảm virut sốt xuất huyết được dự tính trong bản mô tả này có thể được kết hợp với một hoặc nhiều virut sốt xuất huyết sống giảm độc lực. Theo các phương án nhất định, virut sống đã được giảm độc lực có thể là virut DEN-2 sống giảm độc lực trong đó sự đảo ngược thành các axit amin kiểu dại ở các vùng NCR, NS1 hoặc các vùng khác làm gia tăng đáp ứng miễn dịch, làm tăng khả năng sinh trưởng của virut hoặc cải tiến khác đối với cấu trúc virut sốt xuất huyết sống đã được giảm độc lực được cải tiến để dùng trong các chế phẩm được bộc lộ trong sáng chế này.

Theo các phương án nhất định, virut sốt xuất huyết sống giảm độc lực có thể bao gồm đột biến hoặc thay thế đối với nucleotit 5'NCR-57-T, NS1-53-Asp, và NS3-250-Val, của DENV-2, mà có thể tham gia vào cơ sở di truyền đặc hiệu virut PDK-53 của bốn virut cấu trúc TDV. Trình tự gen của ba locus giảm độc lực, cũng như là, các kiểu hình giảm độc lực *in vitro* và *in vivo* đã được thiết lập trước đây của ba ứng

cử viên vacxin này đã được giám sát cẩn thận đối với các giống cấy TDV sản sinh cGMP. Bản mô tả này bộc lộ các chiến lược dùng để tạo ra các chủng virut giống gốc (MVS) cũng như là các giống cấy chuyển tương ứng khác của virut sốt xuất huyết để dùng trong sản xuất chế phẩm vacxin virut sốt xuất huyết. Các MVS này có thể được sử dụng để sản xuất các nguyên liệu lâm sàng và cuối cùng là các nguồn cung cấp vacxin thương mại.

Theo các phương án nhất định, chế phẩm sinh miễn dịch có thể bao gồm các chế phẩm tam giá hoặc tứ giá có khả năng gây ra đáp ứng miễn dịch đối với ít nhất ba hoặc cả bốn typ huyết thanh virut sốt xuất huyết ở trẻ em và thanh niên. Ví dụ, chế phẩm sinh miễn dịch có thể bao gồm các cấu trúc Pre-Master Virut Seed, Master Virut Seed (MVS), Working Virut Seed (WVS) và Bulk Virut Seed (BVS) của mỗi trong số bốn typ huyết thanh virut sốt xuất huyết. Theo một phương án khác, chế phẩm sinh miễn dịch có thể bao gồm một hoặc nhiều polynucleotit có trình tự axit nucleic mã hóa typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực được biểu thị bởi SEQ ID NO: 9 (Pre-Master), 11 (MVS), 13 (WVS) hoặc 15 (BVS); polynucleotit khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2 có trình tự axit nucleic mã hóa protein không cấu trúc từ typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực và protein cấu trúc từ typ huyết thanh virut sốt xuất huyết-1, được biểu thị bởi SEQ ID NO: 1 (Pre-Master), 3 (MVS), 5 (WVS) hoặc 7 (BVS); polynucleotit khảm virut sốt xuất huyết-3/virut sốt xuất huyết-2 có trình tự axit nucleic mã hóa protein không cấu trúc từ typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực và protein cấu trúc từ typ huyết thanh virut sốt xuất huyết-3, được biểu thị bởi SEQ ID NO: 17 (Pre-Master), 19 (MVS), 21 (WVS) hoặc 23 (BVS); và polynucleotit khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 có trình tự axit nucleic mã hóa protein không cấu trúc từ typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực và protein cấu trúc từ typ huyết thanh virut sốt xuất huyết-4 được biểu thị bởi SEQ ID NO: 25 (Pre-Master), 27 (MVS), 29 (WVS) hoặc 31 (BVS). Một số phương án bộc lộ chế phẩm sinh miễn dịch có khả năng gây ra đáp ứng miễn dịch chống lại một, hai, ba hoặc cả bốn typ huyết thanh virut sốt xuất huyết ở trẻ em hoặc thanh niên.

Theo các phương án nhất định, chế phẩm sinh miễn dịch có thể bao gồm các chế phẩm trong đó tỷ lệ nồng độ của polynucleotit mã hóa typ huyết thanh virut sốt xuất huyết-2 sống được cải biến giảm độc lực trong chế phẩm có thể nhỏ hơn ít nhất một nửa đến một log đơn vị hình thành vết tan (plaque forming unit - PFU) log PFU của một hoặc nhiều polynucleotit virut sốt xuất huyết. Theo một phương án khác, chế phẩm sinh miễn dịch có thể bao gồm các chế phẩm trong đó tỷ lệ nồng độ của polynucleotit mã hóa typ huyết thanh virut sốt xuất huyết-2 sống được cải biến giảm độc lực trong chế phẩm ít nhất là nhỏ hơn một và một nửa hoặc lớn hơn hai hoặc ba log đơn vị hình thành vết tan (PFU) log PFU của một hoặc nhiều polynucleotit.

Theo các phương án nhất định, chế phẩm sinh miễn dịch chống lại virut sốt xuất huyết một mình hoặc kết hợp với các vacxin chống virut khác có thể là hỗn hợp gồm các polypeptit và polynucleotit. Theo các phương án này, chế phẩm sinh miễn dịch có thể bao gồm polynucleotit được mô tả ở đây hoặc các polynucleotit flavivirut khác cùng hoặc không cùng với các polypeptit khác nhau được mã hóa bởi các polynucleotit Ví dụ, polypeptit có thể bao gồm polypeptit mà tương ứng với typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực được biểu thị bởi SEQ ID NO: 10 (Pre-Master), 12 (MVS), 14 (WVS) hoặc 16 (BVS); polypeptit tương ứng với polynucleotit khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2 được biểu thị bởi SEQ ID NO: 2 (Pre-Master), 4 (MVS), 6 (WVS) hoặc 8 (BVS); polypeptit tương ứng với polynucleotit khảm virut sốt xuất huyết-3/virut sốt xuất huyết-2 được biểu thị bởi SEQ ID NO: 18 (Pre-Master), 20 (MVS), 22 (WVS) hoặc 24 (BVS); và polypeptit tương ứng với polynucleotit khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 được biểu thị bởi SEQ ID NO: 26 (Pre-Master), 28 (MVS), 30 (WVS) hoặc 32 (BVS).

Theo các phương án nhất định, chế phẩm sinh miễn dịch được bộc lộ ở đây có thể bao gồm polypeptit và/hoặc polynucleotit ở các tỷ lệ nồng độ khác nhau, bao gồm tỷ lệ nồng độ của thể khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2, với typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực, với thể khảm virut sốt xuất huyết-3/virut sốt xuất huyết-2, với thể khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 là 4:4:5:5 log đơn vị hình thành vết tan (PFU). Theo một

phương án khác, tỷ lệ nồng độ của thể khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2, với typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực, với thể khảm virut sốt xuất huyết-3/virut sốt xuất huyết-2, với thể khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 có thể bằng khoảng 5:4:5:5 log PFU. Các tỷ lệ khác được dự tính ở đây có thể bao gồm chế phẩm trong đó thể khảm virut sốt xuất huyết-2 và/hoặc virut sốt xuất huyết-1 sống giảm độc lực là giảm so với ít nhất một trong số thể khảm virut sốt xuất huyết-3 hoặc virut sốt xuất huyết-4 và/hoặc virut sốt xuất huyết-1. Theo một phương án khác nữa, tỷ lệ chế phẩm tứ giá có thể là khoảng 5:3:5:5, 3:3:5:5, hoặc tỷ lệ bất kỳ có trong các ví dụ nhất định làm giảm virut sốt xuất huyết-2 sống giảm độc lực hoặc làm giảm virut điển hình hoặc cả hai virut sốt xuất huyết-2 sống giảm độc lực và thể khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2 trong chế phẩm.

Theo một phương án khác, chế phẩm được bộc lộ ở đây có thể là dược phẩm mà bao gồm một hoặc nhiều chế phẩm sinh miễn dịch được bộc lộ ở đây và tá dược dược dụng. Theo các phương án nhất định, dược phẩm có thể bao gồm chế phẩm sinh miễn dịch được tạo ra từ virut sốt xuất huyết sống giảm độc lực và/hoặc thể khảm virut sốt xuất huyết chứa cả polynucleotit và polypeptit nhằm chống lại một hoặc nhiều typ huyết thanh virut sốt xuất huyết.

Các phương án khác có thể bao gồm kit chứa ít nhất chế phẩm sinh miễn dịch theo sáng chế và ít nhất một đồ chứa dùng để chống lại sự nhiễm virut sốt xuất huyết ở trẻ em và thanh niên.

Các phương án nhất định bao gồm phương pháp gây miễn dịch ở trẻ em và thanh niên đối với ít nhất một đến cả bốn typ huyết thanh virut sốt xuất huyết. Các phương pháp có thể bao gồm bước cho đối tượng sử dụng một hoặc nhiều liều dược phẩm được bộc lộ ở đây. Theo các phương án này, tỷ lệ để có được các đặc tính này có thể bao gồm virut sống đã được giảm độc lực và/hoặc thể khảm mà gây ra đáp ứng miễn dịch đối với cả bốn typ huyết thanh virut sốt xuất huyết ở đối tượng.

Theo các phương án nhất định, phương pháp có thể bao gồm bước cho trẻ em hoặc thanh niên sử dụng chế phẩm sinh miễn dịch bao gồm một hoặc nhiều polypeptit được mã hóa bởi polynucleotit được bộc lộ ở đây ngoài các polynucleotit.

Phương pháp có thể bao gồm các polypeptit mà bao gồm một hoặc nhiều polypeptit được biểu thị bởi SEQ ID NO: 10 (Pre-Master), 12 (MVS), 14 (WVS) hoặc 16 (BVS); polypeptit được biểu thị bởi SEQ ID NO: 2 (Pre-Master), 4 (MVS), 6 (WVS) hoặc 8 (BVS); polypeptit được biểu thị bởi SEQ ID NO: 18 (Pre-Master), 20 (MVS), 22 (WVS) hoặc 24 (BVS); và polypeptit được biểu thị bởi SEQ ID NO: 26 (Pre-Master), 28 (MVS), 30 (WVS) hoặc 32 (BVS). Theo các phương án nhất định, phương pháp có thể bao gồm bước cho đối tượng nằm trong khoảng từ khoảng 1 tuổi đến khoảng 20 tuổi sử dụng chế phẩm sinh miễn dịch chứa cả bốn typ huyết thanh virut sốt xuất huyết được biểu thị trong chế phẩm sinh miễn dịch được bộc lộ ở đây và gây ra đáp ứng miễn dịch đối với cả bốn typ huyết thanh virut sốt xuất huyết (chế phẩm tứ giá) ở đối tượng. Theo một phương án khác, phương pháp có thể bao gồm bước cho đối tượng nằm trong khoảng từ khoảng 1 tuổi đến khoảng 45 tuổi sử dụng chế phẩm sinh miễn dịch được bộc lộ ở đây.

Theo các phương án nhất định, phương pháp có thể bao gồm bước cho đối tượng sử dụng một hoặc nhiều liều chế phẩm tứ giá theo sáng chế theo phương thức bất kỳ chấp nhận được, bao gồm chẳng hạn, dùng dưới da, dùng trong tĩnh mạch, trong chân bì, qua da, qua đường miệng, qua đường xông hít, trong âm đạo, dùng tại chỗ, dùng trong mũi hoặc qua đường trực tràng. Theo các phương án này, chế phẩm được bộc lộ ở đây có thể được dùng ở dạng liều đơn hoặc ở dạng hai hoặc nhiều hơn hai liều cho đối tượng. Theo một số phương án, các chế phẩm sinh miễn dịch được bộc lộ ở đây có thể được dùng trong vòng 90 ngày mỗi loại, trong vòng khoảng 60 ngày mỗi loại, trong vòng khoảng 30 ngày mỗi loại, và trong vòng ít hơn 30 ngày mỗi loại. Theo một số phương án, chế phẩm sinh miễn dịch được bộc lộ ở đây có thể được dùng cho trẻ em hoặc thanh niên ở ngày 0 và dùng lại trong khoảng 3 tháng tính từ lần dùng đầu tiên.

Theo các phương án nhất định, chế phẩm tứ giá được bộc lộ ở đây có thể gây ra đáp ứng miễn dịch đối với cả bốn typ huyết thanh virut sốt xuất huyết ở ít nhất 60% trẻ em và thanh niên nhận một hoặc nhiều liều chế phẩm tứ giá.

Mô tả vắn tắt các hình vẽ

Các hình vẽ dưới đây tạo thành một phần của bản mô tả sáng chế và được đưa vào nhằm minh họa thêm về các phương án nhất định. Một số phương án có thể được hiểu rõ hơn bằng cách tham khảo một hoặc nhiều hình vẽ này ở dạng một mình hoặc kết hợp với phần mô tả chi tiết các phương án cụ thể được thể hiện ở đây.

Fig. 1 là đồ thị được nêu làm ví dụ phản ánh ví dụ về cấu trúc khảm theo sáng chế, DEN-2/DEN-4 so với cấu trúc đã được tạo ra trước đó và virut sốt xuất huyết kiểu dại.

Fig. 2 là biểu đồ tần số được nêu làm ví dụ so sánh các đáp ứng khác nhau bằng cách sử dụng bộ khung DEN-2 sống giảm độc lực (với các đột biến bổ sung) và một typ huyết thanh virut sốt xuất huyết thứ hai làm thành phần cấu tạo thay cho thành phần cấu tạo virut sốt xuất huyết-2 (*chẳng hạn*, TDV-1 MVS). Biểu đồ này minh họa kích thước vết tan của TDV MVS. Virut sốt xuất huyết kiểu dại và các virut ứng viên dùng làm vacxin loại nghiên cứu đã được công bố trước đó được bao gồm để đối chứng và so sánh. Biểu đồ này minh họa sản lượng cấu trúc virut sốt xuất huyết được cải thiện so với cấu trúc khảm virut sốt xuất huyết đối chứng.

Fig. 3 là biểu đồ tần số được nêu làm ví dụ thể hiện tính nhạy với nhiệt độ của TDV MVS (Master Virut Seed, trong bản mô tả này gọi là DENVax). Virut sốt xuất huyết kiểu dại và các virut ứng viên dùng làm vacxin loại nghiên cứu đã được công bố trước đó được bao gồm để so sánh với loại MVS.

Fig. 4 là biểu đồ mà thể hiện sự sinh trưởng của virut của TDV (còn được biết đến như là DENVax) MVS ở tế bào C6/36 so với mẫu đối chứng. Virut sốt xuất huyết kiểu dại và các virut ứng viên dùng làm vacxin loại nghiên cứu được bao gồm để so sánh với TDV MVS (trong bản mô tả này được gọi là DENVax).

Fig. 5A-5C là các đồ thị được nêu làm ví dụ của tính độc thần kinh ở chuột nhắt mới sinh. Kết quả chung của một số thí nghiệm tóm tắt tính độc thần kinh của virut DENV-2 16681 kiểu dại ở chuột nhắt mới sinh CDC-ICR (n=72) và Taconic-ICR (n=32) được cho thử trong hộp sọ với 10^4 pfu virut (A). Tính độc thần kinh của TDV MVS (Master Virut Seed, trong bản mô tả này được gọi là DENVax) được thử

nghiệm ở chuột nhắt Taconic-ICR với liều lượng 10^4 pfu (B) hoặc 10^3 pfu (C). Số động vật được thử nghiệm trên một nhóm ở một thí nghiệm (n=16) hoặc hai thí nghiệm gộp lại (n=31 hoặc 32) được chỉ ra.

Fig. 6 là biểu đồ tần số được nêu làm ví dụ minh họa kích thước vết tan của TDV MVS, WVS, và BVS. Đường kính vết tan trung bình ± SD (khoảng sai số) của vết tan virut ở tế bào Vero hoặc LLC-MK_2 trong điều kiện phủ agaroza được đo ở ngày 9 sau gây nhiễm. Các DENV kiểu dại và các virut ứng viên dùng làm vacxin loại nghiên cứu đã được công bố trước đó được bao gồm để đối chứng và so sánh.

Fig. 7 là biểu đồ tần số được nêu làm ví dụ minh họa sự sinh trưởng của TDV (trong bản mô tả này được gọi là DENVax) MSV, WVS, và BVS ở các tế bào C6/36 ở hai nhiệt độ ủ để xác nhận sự giữ lại của gen đánh dấu giảm độc lực *in vitro* trong chúng sau khi sản xuất ở quy mô lớn. Virut sốt xuất huyết kiểu dại nhất định và các virut ứng viên dùng làm vacxin loại nghiên cứu đã được công bố trước đó được bao gồm để so sánh.

Fig. 8 là biểu đồ được nêu làm ví dụ biểu thị sự sinh trường hạn chế của TDV MVS (Master Virut Seed, trong bản mô tả này được gọi là DENVax), WVS, và BVS ở các tế bào C6/36. Độ chuẩn trung bình ± SD (khoảng sai số) của virut đã được sao chép trong các tế bào C6/36 7 ngày sau gây nhiễm. Virut sốt xuất huyết kiểu dại nhất định và các virut ứng viên dùng làm vacxin loại nghiên cứu đã được công bố trước đó được bao gồm để so sánh.

Fig. 9A-9B là đồ thị được nêu làm ví dụ về dữ liệu của tính độc thần kinh của TDV MVS (Master Virut Seed, trong bản mô tả này được gọi là DENVax) ở chuột nhắt ICR mới sinh. (A) chủng ngừa trong hộp sọ bằng virut ở liều lượng 10^4 PFU. (B) chủng ngừa trong hộp sọ bằng virut ở liều lượng 10^3 PFU.

Fig. 10 là đồ thị được nêu làm ví dụ so sánh virut sốt xuất huyết-2 sống giảm độc lực mới với virut sốt xuất huyết-2 sống giảm độc lực được tạo ra trước đây.

Fig. 11A là sơ đồ thể hiện của virut sốt xuất huyết sống giảm độc lực và cấu trúc khảm virut sốt xuất huyết-virut sốt xuất huyết (được gọi là TDV) theo sáng chế, bao gồm miền protein điển hình và các vị trí đột biến.

Fig. 11B là công thức bào chế điển hình của một chế phẩm tứ giá, theo một phương án của sáng chế.

Fig. 12A là biểu đồ điển hình của một phác đồ liều ở các đối tượng được thực hiện như là một phần của thử nghiệm lâm sàng, theo một phương án của sáng chế.

Fig. 12B là biểu đồ điển hình của một phác đồ liều ở các đối tượng được thực hiện như là một phần của thử nghiệm lâm sàng, theo một phương án của sáng chế.

Fig. 13A là đồ thị ở dạng đường điển hình mô tả kết quả của các thử nghiệm vi trung hòa (microneutralization tests - MNT) trong đó trung hòa các đáp ứng kháng thể sau khi sử dụng chế phẩm tứ giá cho các đối tượng dương tính huyết thanh, theo một phương án của sáng chế.

Fig. 13B là đồ thị ở dạng đường điển hình mô tả kết quả của các thử nghiệm vi trung hòa (microneutralization tests - MNT) trong đó trung hòa các đáp ứng kháng thể sau khi sử dụng chế phẩm tứ giá cho các đối tượng chưa bao giờ tiếp xúc virut sốt xuất huyết, theo một phương án của sáng chế.

Fig. 14 là đồ thị đường điển hình mô tả tỷ lệ phần trăm của đối tượng thử nghiệm lâm sàng mà dương tính huyết thanh đối với cả bốn typ huyết thanh virut sốt xuất huyết ở các thời điểm khác nhau sau khi sử dụng chế phẩm tứ giá, theo một phương án của sáng chế.

Fig. 15 là đồ thị đường điển hình mô tả tỷ lệ phần trăm của đối tượng thử nghiệm lâm sàng mà dương tính huyết thanh đối với cả bốn typ huyết thanh virut sốt xuất huyết ở các thời điểm khác nhau sau khi sử dụng chế phẩm tứ giá, theo một phương án của sáng chế.

Fig. 16 là đồ thị đường điển hình mô tả tỷ lệ phần trăm của đối tượng thử nghiệm lâm sàng dương tính huyết thanh mà chưa từng nhận dược phẩm gồm cả bốn typ huyết thanh virut sốt xuất huyết trước đó ở các thời điểm khác nhau sau khi sử dụng chế phẩm tứ giá, theo một phương án của sáng chế.

Fig. 17 là đồ thị đường điển hình mô tả tỷ lệ phần trăm của đối tượng thử nghiệm lâm sàng dương tính huyết thanh (trẻ em và người lớn) với nhiều typ huyết

thanh sau khi nhận dược phẩm gồm cả bốn typ huyết thanh virut sốt xuất huyết, theo một phương án của sáng chế.

Mô tả chi tiết sáng chế

Phần định nghĩa

Như được sử dụng trong bản mô tả này, "một" có thể có nghĩa một hoặc nhiều hơn một khoản mục.

Như được sử dụng trong bản mô tả này, mô tả "đối tượng" hoặc "các đối tượng" có thể bao gồm, nhưng không giới hạn ở, động vật có vú như người (chẳng hạn từ trẻ em đến người lớn) hoặc các động vật có vú, động vật đã thuần hóa hoặc hoang dã, chẳng hạn chó, mèo, các thú cưng nuôi trong nhà khác (*chẳng hạn*, chuột hamster, chuột lang, chuột nhắt, chuột cống), chồn sương, thỏ, lợn, ngựa, gia súc, sóc chó, loài gặm nhấm hoang dã, hoặc các động vật vườn thú.

Như được sử dụng ở đây, thuật ngữ "thể khảm virut", "virut khảm," "thể khảm flavivirut" và "flavivirut khảm" có thể có nghĩa là cấu trúc gồm một phần gồm trình tự nucleotit của virut sốt xuất huyết-2 và còn bao gồm trình tự nucleotit mà không bắt nguồn từ virut sốt xuất huyết-2 hoặc từ flavivirut khác. "Thể khảm virut sốt xuất huyết" bao gồm ít nhất hai typ huyết thanh virut sốt xuất huyết khác nhau nhưng không bao gồm một flavivirut khác. Vì vậy, ví dụ về các virut sốt xuất huyết hoặc các flavivirut khác có thể bao gồm, nhưng không giới hạn ở, các trình tự từ virut sốt xuất huyết-1, virut sốt xuất huyết-3, virut sốt xuất huyết-4, virut Zika, virut West Nile, virut gây viêm não Nhật Bản, virut gây viêm não St. Louis, virut gây viêm não do tíc gây ra, virut gây sốt vàng da và tổ hợp bất kỳ của chúng.

Như được sử dụng ở đây, "thể khảm axit nucleic" có thể có nghĩa cấu trúc theo sáng chế bao gồm axit nucleic mang một phần trình tự nucleotit của virut sốt xuất huyết (chẳng hạn virut sốt xuất huyết-2) và còn bao gồm các trình tự nucleotit mà không có cùng nguồn gốc với virut sốt xuất huyết. Theo các phương án này, virut sốt xuất huyết khảm hoặc thể khảm flavivirut bất kỳ được bộc lộ ở đây có thể được thừa nhận như là ví dụ về thể khảm axit nucleic.

Như được sử dụng ở đây, "virut sống đã được giảm độc lực" có thể có nghĩa

là virut kiểu dại, được đột biến hoặc chọn lọc các tính trạng để dùng trong vacxin hoặc các chế phẩm sinh miễn dịch khác trong đó một số tính trạng có thể bao gồm tính độc giảm, độ an toàn, hiệu lực hoặc khả năng sinh trưởng gia tăng v.v.

Như được sử dụng ở đây, "chế phẩm virut sốt xuất huyết" hoặc "(các)chế phẩm sinh miễn dịch" hoặc "chế phẩm vacxin" hoặc dược phẩm chứa nó có thể bao gồm các tổ hợp khác nhau gồm các polynucleotit và/hoặc polypeptit được bộc lộ ở đây, mà có thể được dùng cho đối tượng và có thể gây ra đáp ứng miễn dịch cho một hoặc nhiều typ huyết thanh virut sốt xuất huyết ở đối tượng đó.

Phần mô tả

Trong phần sau đây, ví dụ về các chế phẩm và phương pháp khác nhau được mô tả để cụ thể hóa các phương án khác nhau. Chuyên gia trong lĩnh vực hiểu rõ rằng việc thực hiện các phương án khác nhau không đòi hỏi sử dụng tất cả hoặc thậm chí là một vài chi tiết cụ thể được nêu ra ở đây, mà thay vào đó nồng độ, thời gian và các chi tiết cụ thể khác có thể được cải biến nhờ thí nghiệm thông thường. Trong một số trường hợp, các phương pháp hoặc thành phần đã biết không được bao hàm trong phần mô tả.

Theo các phương án của sáng chế, các kỹ thuật sinh học phân tử, hóa protein, vi sinh, và ADN tái tổ hợp thông thường có thể được sử dụng bởi chuyên gia trong lĩnh vực. Các kỹ thuật như vậy được giải thích đầy đủ trong tài liệu. Ví dụ, xem tài liệu Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Second Edition 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Animal Cell Culture, R. I. Freshney, ed., 1986).

Các phương án được bộc lộ ở đây liên quan đến chế phẩm, phương pháp, và các ứng dụng để gây ra đáp ứng miễn dịch đối với một hoặc lên tới cả bốn typ huyết thanh virut sốt xuất huyết ở đối tượng như trẻ em hoặc thanh niên từ khoảng 1 tuổi đến khoảng 20 tuổi. Các phương án khác liên quan đến chế phẩm, phương pháp và các ứng dụng để gây ra đáp ứng miễn dịch đối với một hoặc nhiều typ huyết thanh virut sốt xuất huyết ở đối tượng, riêng biệt hoặc đồng thời. Theo các phương án này, virut sốt xuất huyết giảm độc lực và thể khảm axit nucleic có thể được tạo ra và được

sử dụng trong chế phẩm vacxin được bộc lộ ở đây. Một số phương án liên quan đến cấu trúc hoặc thể khảm virut sốt xuất huyết được cải biến hoặc đột biến. Các phương án khác liên quan đến việc đưa các đột biến vào để cải biến trình tự axit amin của protein cấu trúc của virut sốt xuất huyết trong đó đột biến này làm tăng tính sinh miễn dịch đối với virut này.

Virut sốt xuất huyết sống giảm độc lực thuộc cả bốn typ huyết thanh đã được phát triển bằng cách cấy chuyển các virut kiểu dại trong môi trường nuôi cấy tế bào. Các virut này là một vài trong số các ứng viên vacxin sống giảm độc lực triển vọng nhất để chủng ngừa chống lại flavivirut và cụ thể là sự nhiễm virut sốt xuất huyết và/hoặc bệnh sốt xuất huyết. Các ứng viên vacxin này được gọi tên theo tổ hợp gồm typ huyết thanh virut sốt xuất huyết của chúng, dòng tế bào mà thông qua đó chúng được cấy chuyển và số lần chúng được cấy chuyển. Vì vậy, virut sốt xuất huyết kiểu dại typ huyết thanh 1 được cấy chuyển trong tế bào PDK 13 lần được gọi là virut DEN-1 PDK-13. Các ứng viên vacxin khác là DEN-2 PDK-53, DEN-3 PGMK-30/FRhL-3 (*chẳng hạn*, ba mươi lần cấy chuyển ở tế bào thận sơ cấp ở khỉ mặt xanh, sau đó là ba lần cấy chuyển ở tế bào phổi bào thai khỉ nâu Ấn Độ và DEN-4 PDK-48). Bốn virut vacxin ứng viên này được bắt nguồn từ quá trình cấy chuyển dịch nuôi cấy mô của các virut DEN-1 16007, DEN-2 16681, DEN-3 16562 và DEN-4 1036 chủng bố mẹ kiểu dại, tương ứng.

Ứng viên vacxin virut DEN-2 PDK-53 (dưới đây gọi là PDK-53) có một vài dấu chuẩn sinh học có thể đo được có liên quan đến việc giảm độc lực, bao gồm độ nhạy với nhiệt độ, kích thước vết tan nhỏ, sao chép giảm ở dịch nuôi cấy tế bào C6136 của muỗi, sao chép giảm ở muỗi nguyên vẹn, mất tính độc thần kinh đối với chuột còn đang bú và giảm tỷ lệ mắc phải sự nhiễm trùng máu ở khỉ. Thử nghiệm lâm sàng của vacxin ứng viên PDK-53 đã chứng minh độ an toàn và tính sinh miễn dịch của nó trên người. Ngoài ra, vacxin PDK-53 gây ra đáp ứng ghi nhớ của tế bào T đặc hiệu virut sốt xuất huyết ở người nhận vacxin. Một số phương án ở đây mô tả DEN-2 PDK-53 cải tiến sử dụng trong các cấu trúc khảm được bộc lộ ở đây để gây ra đáp ứng miễn dịch đối với virut sốt xuất huyết ở trẻ em và thanh niên.

Thể khảm flavivirut sinh miễn dịch mang bộ khung virut sốt xuất huyết-2 và

ít nhất một protein cấu trúc của một typ huyết thanh virut sốt xuất huyết khác có thể được sử dụng để điều chế thể khảm virut sốt xuất huyết và phương pháp sản xuất thể khảm virut sốt xuất huyết được mô tả. Thể khảm virut sốt xuất huyết sinh miễn dịch được đề xuất, một mình hoặc ở dạng kết hợp, trong chất mang dược dụng để làm chế phẩm sinh miễn dịch để giảm thiểu, ức chế, hoặc chủng ngừa cho những cá nhân chống lại sự nhiễm một hoặc nhiều typ huyết thanh, như các typ huyết thanh virut sốt xuất huyết DEN-1, DEN-2, DEN-3 and DEN-4, một mình hoặc kết hợp. Khi được kết hợp, thể khảm virut sốt xuất huyết sinh miễn dịch có thể được sử dụng để làm vacxin đa giá (*chẳng hạn*, nhị, tam và tứ giá) để có được sự bảo vệ đồng thời đối với sự nhiễm bởi nhiều hơn một loài hoặc chủng flavivirut. Theo các phương án nhất định, thể khảm virut sốt xuất huyết có thể được kết hợp trong chế phẩm sinh miễn dịch để dùng làm vacxin nhị giá, tam giá hoặc tứ giá chống lại các typ huyết thanh virut sốt xuất huyết đã biết có thể có tính miễn dịch với các flavivirut gây bệnh khác do chứa các axit nucleic mã hóa một hoặc nhiều protein từ flavivirut khác.

Theo một số phương án, thể khảm virut sốt xuất huyết sinh miễn dịch không có độc tính được nêu trong bản mô tả chứa các gen protein không cấu trúc của virut sốt xuất huyết-2 giảm độc lực (*chẳng hạn*, PDK-53), hoặc dạng tương đương của nó, và một hoặc nhiều gen protein cấu trúc hoặc phần sinh miễn dịch của nó của flavivirut mà gây ra tính sinh miễn dịch chống lại nó ở đối tượng. Ví dụ, một số phương án liên quan đến thể khảm chứa bộ gen virut sốt xuất huyết-2 đã được giảm độc lực PDK-53 làm bộ khung virut, và một hoặc nhiều gen protein cấu trúc mã hóa lớp vỏ capsit, tiền màng/màng, hoặc vỏ ngoài của bộ gen PDK-53, hoặc tổ hợp của chúng, được thay bằng một hoặc nhiều gen protein cấu trúc tương ứng từ DEN-1, DEN-3 hoặc DEN-4 hoặc flavivirut khác để được bảo vệ chống lại, chẳng hạn một flavivirut khác hoặc một typ huyết thanh virut sốt xuất huyết khác. Theo các phương án này, thể khảm axit nucleic được bộc lộ ở đây có thể có các đặc tính chức năng của virut sốt xuất huyết-2 giảm độc lực và không độc, nhưng biểu hiện epitop kháng nguyên của sản phẩm gen cấu trúc của DEN-1, DEN-3 hoặc DEN-4 ngoài các flavivirut khác và sinh miễn dịch (*chẳng hạn*, gây ra đáp ứng miễn dịch đối với các sản phẩm gen ở đối tượng). Tiếp theo, các cấu trúc này được cấu chuyển 1 hoặc

nhiều lần đến tận 53 lần cấy chuyển để sản sinh chế phẩm virut sống đã được giảm độc lực dùng để tạo ra chế phẩm sinh miễn dịch đối với một hoặc nhiều typ huyết thanh virut sốt xuất huyết (*chẳng hạn*, P1-P10).

Theo một phương án khác, thể khảm axit nucleic có thể là thể khảm axit nucleic mang, nhưng không giới hạn ở, trình tự nucleotit thứ nhất mã hóa protein không cấu trúc từ virut sốt xuất huyết-2 giảm độc lực, và trình tự nucleotit thứ hai mã hóa protein cấu trúc từ virut sốt xuất huyết-4 một mình hoặc kết hợp với một flavivirut khác. Theo một phương án khác, virut sốt xuất huyết-2 giảm độc lực có thể là chủng PDK-53 dùng cho vacxin chứa một hoặc nhiều axit amin đã được đột biến ngược so với axit amin kiểu dại (xem phần ví dụ) được lựa chọn để có các đặc tính vật lý được cải thiện hoặc tính sinh miễn dịch gia tăng. Các đột biến ngược đặc hiệu này trao cho virut các tính trạng mong muốn để dùng làm virut sốt xuất huyết-2 sống giảm độc lực hoặc cấu trúc khảm được mô tả ở đây trong các phác đồ chống lại sự nhiễm virut sốt xuất huyết ở trẻ em và thanh niên. Một số phương án bao gồm protein cấu trúc của một hoặc nhiều trong số các protein C, prM hoặc E của virut sốt xuất huyết thứ hai.

Theo một phương án khác được bộc lộ ở đây, nucleotit và các trình tự axit amin có thể bao gồm các đoạn thay thế, đoạn khuyết hoặc cài xen chẳng hạn, trong bộ gen virut sốt xuất huyết-2 PDK-53 để làm giảm sự can thiệp của các typ huyết thanh virut sốt xuất huyết mục tiêu khác. Các cải biến này có thể được tạo ra trong protein cấu trúc và không cấu trúc một mình hoặc kết hợp với các cải biến được nêu làm ví dụ được bộc lộ ở đây và có thể được tạo ra bằng cách cấy chuyển virut giảm độc lực và thu được chế phẩm cải tiến để gây ra đáp ứng miễn dịch đối với một hoặc nhiều typ huyết thanh virut sốt xuất huyết.

Các phương án nhất định được bộc lộ ở đây đề xuất phương pháp tạo ra flavivirut khảm bằng cách sử dụng kỹ thuật tái tổ hợp, chẳng hạn, bằng cách cài xen trình tự thay thế của một typ huyết thanh vào một bộ khung typ huyết thanh khác. Các phương án khác trong bản mô tả này liên quan đến việc cấy chuyển một virut khảm sống giảm độc lực (*chẳng hạn*, an toàn và hiệu quả) để cải thiện hơn nữa và chọn lọc các tính trạng mong muốn. Theo các phương án nhất định, virut sốt xuất

huyết-2 sống giảm độc lực được sử dụng ở đây có thể bao gồm một hoặc nhiều đột biến được nêu trong bảng 3. Theo một phương án khác, thể khảm virut sốt xuất huyết-virut sốt xuất huyết theo sáng chế có thể bao gồm một hoặc nhiều đột biến như được thể hiện trong bảng 3. Theo các phương án khác nữa, thể khảm virut sốt xuất huyết-virut sốt xuất huyết có thể bao gồm tất cả các đột biến đối với mỗi thể khảm như được thể hiện trong bảng 3 đối với Den-2/Den-1, Den-2/Den-3 hoặc Den-2/Den-4. Dược phẩm bao gồm virut sống đã được giảm độc lực được biểu thị bởi các cấu trúc trong bảng 3 được dự tính. Chẳng hạn, chế phẩm đơn giá, nhị giá, tam giá hoặc tứ giá được dự tính để dùng ở đây bằng cách sử dụng thể khảm virut sốt xuất huyết-virut sốt xuất huyết và virut sốt xuất huyết-2 sống giảm độc lực như được thể hiện trong bảng 3.

Theo các phương án nhất định, biến thể DEN-2 sống giảm độc lực được dự tính ở đây có thể được bào chế thành dược phẩm trong đó dược phẩm có thể được dùng một mình hoặc kết hợp với thể khảm virut sốt xuất huyết-virut sốt xuất huyết hoặc thể khảm virut sốt xuất huyết-flavivirut. Theo các phương án nhất định, chế phẩm nhị giá, tam giá hoặc tứ giá có thể được dùng cho đối tượng trong một ứng dụng duy nhất hoặc trong nhiều ứng dụng.

Thể khảm flavivirut

Virut sốt xuất huyết typ 1-4 (DEN-1 đến DEN-4) là nguồn bệnh flavivirut được sinh ra ở muỗi. Bộ gen flavivirut chứa vùng không mã hóa 5' (5'-NC), vùng mã hóa protein capsit (C), vùng mã hóa protein tiền màng/màng (prM), vùng mã hóa protein vỏ ngoài (E), vùng mã hóa protein không cấu trúc (NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5) và vùng không mã hóa 3' (3'NC). Protein cấu trúc flavivirut bao gồm C, prM và E. Protein không cấu trúc bao gồm NS1-NS5. Protein cấu trúc và không cấu trúc được dịch mã thành một polyprotein đơn và được xử lý bởi proteaza tế bào và virut.

Thể khảm flavivirut có thể là các cấu trúc được tạo thành bằng cách dung hợp các gen protein không cấu trúc từ một loại, hoặc typ huyết thanh, của virut sốt xuất huyết hoặc các loài virut của flaviviridae, với các gen protein, chẳng hạn, các gen protein cấu trúc, từ một loại, hoặc typ huyết thanh, khác của virut sốt xuất huyết hoặc

flavivirut khác. Theo một phương án khác, cấu trúc khảm flavivirut có thể được tạo thành bằng cách dung hợp các gen protein không cấu trúc từ một loại, hoặc typ huyết thanh, của virut sốt xuất huyết hoặc các loài virut của flaviviridae, với các trình tự nucleotit khác mà điều khiển quá trình tổng hợp polypeptit hoặc protein được lựa chọn từ các typ huyết thanh virut sốt xuất huyết khác hoặc các flavirut khác. Theo một phương án khác, việc thế các trình tự của một flavivirut cho một flavivirut khác được dự tính.

Theo một phương án khác, thể khảm virut sốt xuất huyết có thể chứa các gen protein không cấu trúc của virut sốt xuất huyết sống đã được giảm độc lực, hoặc dạng tương đương của nó, và một hoặc nhiều gen protein cấu trúc, hoặc các phần kháng nguyên của nó, của virut sốt xuất huyết hoặc flavivirut khác mà có tính sinh miễn dịch chống lại chúng.

Các virut sốt xuất huyết thích hợp khác để dùng trong xây dựng thể khảm virut sốt xuất huyết có thể là kiểu dại, có tính độc DEN-1 16007, DEN-2 16681, DEN-3 16562 và DEN-4 1036 và chủng vacxin giảm độc lực DEN-1 PDK-13, DEN-2 PDK-53, DEN-3 PMK-30/FRhL-3 và DEN-4 PDK-48. Sự khác biệt về mặt di truyền giữa các cặp virut kiểu dại/giảm độc lực DEN-1, DEN-2, DEN-3 và DEN-4 được dự tính cùng với sự thay đổi về trình tự axit amin được mã hóa bởi bộ gen virut.

Theo một số phương án, virut sốt xuất huyết-2 để dùng trong bản mô tả này có thể bao gồm biến thể DEN-2 PDK-53-V, trong đó vị trí nucleotit 5270 trong bộ gen được đột biến từ A thành T và vị trí axit amin 1725 của polyprotein hoặc vị trí axit amin 250 của protein NS3 chứa gốc valin của trình tự virut sốt xuất huyết-2 với chiều dài đầy đủ. Một biến thể DEN-2 PDK-53 không chứa đột biến nucleotit này, DEN-2 PDK-53-E, khác với PDK-53-V ở vị trí này. DEN-2 PDK-53-E chứa A ở vị trí nucleotit 5270 và glutamat ở vị trí polypeptit 1725, protein NS3 ở vị trí axit amin 250. Theo một số phương án, virut sốt xuất huyết-2 PDK 53 cải biến có thể bao gồm một hoặc nhiều đột biến ngược của các vị trí này thành trình tự tự nhiên để có được nhiều tính trạng mong muốn hơn để dùng trong chế phẩm vacxin được dự tính ở đây.

Theo các phương án nhất định, thể khảm virut sốt xuất huyết-virut sốt xuất huyết có thể bao gồm dòng lây nhiễm đặc hiệu virut DEN-2 có bộ khung cải biến và

gen cấu trúc (prM-E hoặc C-prM-E) được cài xen từ các virut sốt xuất huyết khác hoặc các flavivirut khác. Theo một số phương án, biến thể bộ khung virut sốt xuất huyết-2 có thể được tạo ra từ virut sốt xuất huyết-2 16681 chủng (P), PDK-53-E (E), hoặc PDK-53-V (V); chữ cái cuối cùng chỉ ra gen cấu trúc C-prM-E từ chủng bổ mẹ (P) hoặc dẫn xuất vacxin của nó (V) hoặc gen cấu trúc prM-E từ bố mẹ (P) hoặc dẫn xuất vacxin của nó (V1). Chẳng hạn; DEN-2/1-VP biểu thị thể khảm chứa bộ khung DEN-2 PDK-53V giảm độc lực chứa một valin ở NS3-250 và các gen C-prM-E từ DEN-1 16007 kiểu dại; DEN-2/1-VV biểu thị bộ khung DEN-2 PDK-53V với chủng vacxin của virut sốt xuất huyết-1, DEN-1 PDK-13; DEN-2/1-VP1 biểu thị bộ khung DEN-2 PDK-53V và các gen prM-E từ DEN-1 16007 kiểu dại; DEN-2/3-VP1 biểu thị bộ khung DEN-2 PDK-53V và các gen prM-E từ DEN-3 16562 kiểu dại; DEN-2/4VP1 biểu thị bộ khung DEN-2 PDK-53V và các gen prM-E từ DEN-4 1036 kiểu dại. Các thể khảm khác được bộc lộ ở đây được nhận diện bởi cách gọi tên theo danh pháp tương tự.

Theo một số phương án, thể khảm virut sốt xuất huyết-virut sốt xuất huyết được bộc lộ ở đây có thể chứa bộ gen virut sốt xuất huyết-2 giảm độc lực PDK-53 làm bộ khung ở đó các gen protein cấu trúc mã hóa các protein C, prM và E của bộ gen PDK-53, hoặc tổ hợp của chúng, có thể được thay bằng các gen protein cấu trúc tương ứng từ virut sốt xuất huyết-1, virut sốt xuất huyết-3 hoặc virut sốt xuất huyết-4 và tùy ý, một flavivirut khác để bảo vệ chống lại, chẳng hạn như một flavivirut khác (chẳng hạn Zika hoặc sốt vàng da hoặc tổ hợp của chúng) hoặc một chủng virut sốt xuất huyết khác. Chủng virut sốt xuất huyết-2 sống giảm độc lực PDK-53 có kiểu gen kết hợp ở vị trí nucleotit 5270. Một phần đáng kể (khoảng 29%) quần thể virut mã hóa NS3-250-Glu chưa đột biến mà có mặt trong virut DEN-2 16681 kiểu dại hơn là mã hóa thể đột biến NS3-250-Val. Do cả hai thể đột biến di truyền đều không có tính độc, sự đột biến này có thể không cần thiết trong thể khảm không độc.

Theo các phương án nhất định, đột biến đơn ở NS1-53, đột biến kép ở NS1-53 và ở 5'NC-57, đột biến kép ở NS1-53 và ở NS3-250 và đột biến ba ở NS1-53, ở 5'NC-57 và ở NS3-250, làm giảm độc lực của virut DEN-2. Vì vậy, bộ gen của virut sốt xuất huyết-2 bất kỳ chứa các phần tử thế axit amin không bảo toàn hoặc các phần

tử thế nucleotit như vậy ở các locus này có thể được dùng làm trình tự cơ sở để thu được virut PDK-53 cải biến được bộc lộ ở đây. Đột biến khác trong thân của cấu trúc thân/vòng ở vùng không mã hóa 5' có thể còn tạo ra độ ổn định kiểu hình không độc, nếu muốn. Đột biến ở vùng này phá vỡ cấu trúc bậc hai tiềm năng quan trọng đối với sự sao chép virut.

Đột biến được bộc lộ ở đây có thể đạt được bằng phương pháp bất kỳ đã biết trong lĩnh vực bao gồm, nhưng không giới hạn ở, các dòng chọn lọc mang các đặc trưng khác khi được cấy chuyển trong dòng tế bào quan tâm (*chẳng hạn,* tế bào Vero). Chuyên gia trong lĩnh vực hiểu rõ rằng thử nghiệm sàng lọc tính độc, như được mô tả ở đây và đã được biết rõ trong lĩnh vực, có thể được sử dụng để phân biệt giữa các cấu trúc bộ khung có tính độc và không có tính độc.

Xây dựng thể khảm flavivirut

Thể khảm flavivirut được bộc lộ ở đây có thể được sản xuất bằng cách nối một hoặc nhiều gen protein cấu trúc của một flavivirut mà mong muốn có được tính miễn dịch chống lại nó vào bộ khung là bộ gen virut sốt xuất huyết-2 PDK-53, hoặc các phương pháp khác đã biết trong lĩnh vực, bằng cách sử dụng kỹ thuật tái tổ hợp để loại bỏ gen PDK-53 tương ứng và thay thế nó bằng gen virut sốt xuất huyết-1, virut sốt xuất huyết-3 hoặc virut sốt xuất huyết-4 hoặc gen khác đã biết trong lĩnh vực.

Theo cách khác, bằng cách sử dụng các trình tự được nêu trong danh mục trình tự, phân tử axit nucleic mã hóa protein flavivirut có thể được tổng hợp bằng cách sử dụng các kỹ thuật tổng hợp axit nucleic đã biết và cài xen vào vector thích hợp. Virut sống đã được giảm độc lực được bộc lộ ở đây có thể được tạo ra bằng cách sử dụng các kỹ thuật tái tổ hợp mà chuyên gia trong lĩnh vực kỹ thuật này đã biết rõ.

Các virut thể khảm hoặc thể khảm axit nucleic thích hợp chứa trình tự nucleotit mã hóa protein cấu trúc của các flavivirut hoặc typ huyết thanh virut sốt xuất huyết khác có thể được đánh giá tính hữu dụng để làm vacxin bằng cách sàng lọc chúng theo các dấu chuẩn kiểu hình nêu trên về sự giảm độc lực cho thấy không

có tính độc và bằng cách sàng lọc chúng theo tính sinh miễn dịch. Tính kháng nguyên và tính sinh miễn dịch có thể được đánh giá bằng cách sử dụng khả năng phản ứng in vitro hoặc in vivo với các kháng thể flavivirut hoặc huyết thanh phản ứng miễn dịch bằng cách sử dụng các quy trình sàng lọc thông thường mà chuyên gia trong lĩnh vực đã biết rõ.

Vacxin virut sốt xuất huyết

Theo các phương án nhất định, virut thể khảm và thể khảm axit nucleic có thể tạo ra virut sống đã được giảm độc lực hữu hiệu để dùng làm chất sinh miễn dịch hoặc vacxin. Một số phương án bao gồm thể khảm mà thể hiện tính sinh miễn dịch tốt đối với virut sốt xuất huyết-4 trong khi không tạo ra tác dụng gây bệnh nguy hiểm hoặc chết người.

Để làm giảm sự xảy ra DHF/DSS ở trẻ em và thanh niên đã được chủng ngừa đối với chỉ một typ huyết thanh của virut sốt xuất huyết, vacxin tứ giá là cần thiết để tạo ra tính miễn dịch đồng thời đối với cả bốn typ huyết thanh của virut này. Vacxin tứ giá có thể được sản xuất bằng cách kết hợp một virut sốt xuất huyết-2 sống giảm độc lực theo sáng chế với các thể khảm virut sốt xuất huyết-2/1, virut sốt xuất huyết-2/3, and virut sốt xuất huyết-2/4 hoặc một cấu trúc virut sốt xuất huyết khác được mô tả ở đây trong chất mang dược dụng thích hợp để dùng làm vacxin đa giá.

Theo các phương án nhất định, virut thể khảm hoặc thể khảm axit nucleic theo sáng chế có thể bao gồm gen cấu trúc của virut kiểu dại hoặc virut sống đã được giảm độc lực trên bộ khung virut DEN-2 giảm độc lực. Chẳng hạn, thể khảm virut sốt xuất huyết-2/virut sốt xuất huyết-4 có thể biểu hiện các gen protein cấu trúc của virut DEN-4 1036 kiểu dại.

Theo các phương pháp nhất định, virut được sử dụng trong thể khảm được mô tả ở đây có thể được sinh trưởng bằng cách sử dụng các kỹ thuật đã biết trong lĩnh vực. Sau đó, chuẩn độ vết tan virut có thể được thực hiện và vết tan được đếm để đánh giá khả năng sống sót vaf các đặc trưng kiểu hình của giống cấy đang sinh trưởng. Virut kiểu dại có thể được cấy chuyển qua các dòng tế bào đã được nuôi cấy để tạo ra nguyên liệu đầu của virut ứng viên giảm độc lực.

Theo các phương án nhất định, các dòng khảm có thể được xây dựng từ các dòng typ huyết thanh virut sốt xuất huyết khác nhau sẵn có đối với chuyên gia trong lĩnh vực. Việc tạo dòng các mảnh ADN bổ trợ đặc hiệu virut cũng có thể được thực hiện. Các mảnh ADN bổ trợ chứa các gen protein cấu trúc hoặc các gen protein không cấu trúc có thể được khuếch đại bằng phản ứng chuỗi polymeraza phiên mã ngược (reverse transcriptase-polymerase chain reaction - RT-PCR) từ ARN virut sốt xuất huyết với các đoạn mồi khác nhau. Các mảnh đã khuếch đại có thể được tạo dòng vào các vị trí phân cắt của các dòng trung gian khác. Tiếp theo, các dòng virut sốt xuất huyết khảm trung gian có thể được đọc trình tự để kiểm tra độ chính xác của ADN bổ trợ đặc hiệu virut sốt xuất huyết được cài xen.

Các plasmit khảm thuộc bộ gen có chiều dài đầy đủ mà được xây dựng bằng cách cài xen vùng gen protein cấu trúc và/hoặc protein không cấu trúc của virut typ huyết thanh virut sốt xuất huyết vào các vector có thể thu được bằng cách sử dụng kỹ thuật tái tổ hợp mà chuyên gia trong lĩnh vực đã biết rõ.

Phân tích nucleotit và axit amin

Đột biến NS1-53 ở virut vacxin DEN-2 PDK-53 là có ý nghĩa đối với kiểu hình giảm độc lực của virut này, do NS1-53-Gly của virut DEN-2 16681 được bảo toàn ở hầu như tất cả các flavivirut, bao gồm virut sinh ra do tíc, được lập trình tự cho đến nay. Virut vacxin DEN-4 cũng có thể chứa đột biến axit amin ở protein NS1 ở vị trí 253. Locus này, mà là đột biến Gln-to-His ở virut vacxin DEN-4 PDK-48, là Gln ở cả bốn typ huyết thanh kiểu dại của virut sốt xuất huyết. Gốc Gln này chỉ liên quan đến virut sốt xuất huyết thuộc giống flavivirut. Protein NS1 là glycoprotein được tiết từ các tế bào được cấy nhiễm flavivirut. Nó có mặt trên bề mặt của tế bào đã nhiễm và các kháng thể đặc hiệu NS1 có mặt trong huyết thanh của các cá thể bị nhiễm virut. Bảo vệ động vật đã được gây miễn dịch với protein NS1 hoặc thụ động với kháng thể đặc hiệu NS1 đã được báo cáo. Protein NS1 dường như tham gia vào sự sao chép ARN virut ở pha sớm.

Theo các phương án nhất định, đột biến mà xuất hiện trong các protein NS2A, NS2B, NS4A, và NS4B của các chủng DEN-1, -2, -3 và -4 giảm độc lực là các đột biến có tính bảo toàn. Đột biến NS4A-75 và NS4A-95 của các virut vacxin DEN-2

và DEN-4, tương ứng, xuất hiện ở các vị trí của axit amin bảo toàn trong số các virut sốt xuất huyết, nhưng nhìn chung không trong số các flavivirut.

Các trình tự axit nucleic mã hóa virut DEN-4, DEN-3 hoặc DEN-1 (*chẳng hạn*, các phần tử cấu tạo) có thể được cài xen vào một vector, như plasmit, và được biểu hiện tái tổ hợp trong cơ thể sống (*chẳng hạn*, được cài xen vào bộ khung virut sốt xuất huyết-2) để tạo ra các peptit và/hoặc polypeptit của virut sốt xuất huyết tái tổ hợp và/hoặc các virut.

Phương pháp phát hiện axit nucleic

Thử nghiệm di truyền nhanh để chuẩn đoán mỗi trong số các virut vacxin được mô tả ở đây được đề xuất bởi sáng chế này. Phương án này của sáng chế làm nổi bật phép phân tích virut được phân lập từ huyết thanh của người đã được chủng ngừa mà bị nhiễm trùng máu, cũng như làm nổi bật các đặc tính của bệnh nhiễm trùng máu ở động vật linh trưởng không phải người được gây miễn dịch bằng virut vacxin ứng cử.

Các trình tự này có thể bao gồm đoạn dò chẩn đoán TaqMan đóng vai trò thông báo sự phát hiện của đơn vị siêu sao chép ADN bổ trợ được khuếch đại từ khuôn mẫu ARN bộ gen virut bằng cách sử dụng phản ứng chuỗi enzym phiên mã ngược/polymeraza (reverse-transciptase/polymerase chain reaction - RT/PCR), cũng như các các đơn vị khuếch đại xuôi và ngược được thiết kế để khuếch đại các đơn vị siêu sao chép ADN bổ trợ, như được mô tả dưới đây. Trong các trường hợp nhất định, một trong số các đơn vị khuếch đại đã được thiết kế để chứa đột biến đặc hiệu virut vacxin ở đầu tận cùng 3' của đơn vị khuếch đại này, mà nó thực sự khiến cho thử nghiệm này thậm chí đặc hiệu hơn đối với chủng vacxin do sự kéo dài của đoạn mồi ở vị trí đích, và do đó sự khuếch đại sẽ chỉ xảy ra khi khuôn mẫu ARN virut chứa đột biến đặc hiệu này.

Hệ thống phát hiện trình tự axit nucleic tự động trên cơ sở PCR có thể được sử dụng, hoặc kỹ thuật đã biết khác để phát hiện axit nucleic. Thử nghiệm TaqMan là thử nghiệm cực kỳ đặc hiệu và nhạy, cho phép hình dung và định lượng tự động theo thời gian thực các đơn vị siêu sao chép được tạo ra bởi PCR từ khuôn axit

nucleic được dùng làm mẫu. TaqMan có thể xác định sự có mặt hoặc vắng mặt của một trình tự cụ thể. Trong thử nghiệm này, đoạn mồi xuôi và đoạn mồi ngược được thiết kế để gắn ngược chiều và xuôi chiều vị trí đột biến đích, tương ứng. Đoạn dò phát hiện đặc hiệu, mà được thiết kế để có nhiệt độ nóng chảy cao hơn khoảng 10 độ C so với các đơn vị khuếch đại và chứa đột biến nucleotit đặc hiệu virut vacxin hoặc phần bổ trợ của nó (tùy thuộc vào sợi đơn vị siêu sao chép RT/PCR được phát hiện), tạo thành thành phần mồi thứ ba của thử nghiệm.

Đoạn dò được thiết kế để phát hiện đặc hiệu locus đột biến ở một trong số các bộ gen virut vacxin sẽ chứa nucleotit đặc hiệu vacxin ở giữa của đoạn dò. Đoạn dò này sẽ tạo ra sự phát huỳnh quang có thể phát hiện được trong thử nghiệm TaqMan nếu khuôn mẫu ARN virut là đặc hiệu đối với virut dùng làm vacxin. Tuy nhiên, khuôn mẫu ARN bộ gen từ các virut DEN kiểu dại sẽ có hiệu quả lai hóa đầu dò giảm đáng kể do sự không ghép đôi của một nucleotit đơn (trong trường hợp các virut DEN là virut bố mẹ) hoặc có thể không ghép đôi của nhiều hơn một nucleotit (như có thể xảy ra trong các virut DEN kiểu dại khác) và sẽ không tạo ra sự phát huỳnh quang đáng kể. ADN polymeraza có khả năng lớn hơn trong việc thay thế đoạn dò không ghép đôi từ khuôn khuếch đại RT/PCR so với tách đoạn dò không ghép đôi để giải phóng thuốc nhuộm thông báo (thử nghiệm TaqMan Allelic Discrimination, Applied Biosystems).

Một chiến lược để nghiên cứu chẩn đoán di truyền là sử dụng các đèn báo hiệu phân tử. Chiến lược báo hiệu phân tử này cũng sử dụng các đoạn mồi để khuếch đại các đơn vị siêu sao chép nhờ RT/PCR, và phát hiện trình tự cụ thể trong đơn vị siêu sao chép này bằng một đầu dò chứa thuốc nhuộm thông báo và dập tắt ở các đầu tận cùng của đầu dò. Trong thử nghiệm này, đầu dò tạo thành cấu trúc thân-vòng. Thử nghiệm báo hiệu phân tử sử dụng các thuốc nhuộm dập tắt và thông báo khác với các thuốc nhuộm được sử dụng trong thử nghiệm TaqMan.

Dược phẩm

Các phương án trong bản mô tả này đề xuất việc sử dụng chế phẩm ở dạng tương thích về mặt sinh học thích hợp để dùng *in vivo* trong dược phẩm cho các đối tượng. Thuật ngữ "dạng tương thích về mặt sinh học thích hợp để dùng in vivo" có

nghĩa là dạng hoạt chất (*chẳng hạn*, dược chất, theo phương án này) để dùng trong đó tác dụng trị liệu của hoạt chất lớn hơn tác dụng gây độc bất kỳ của hoạt chất này. Việc dùng một lượng hữu hiệu có tác dụng trị liệu của chế phẩm trị liệu được định nghĩa là lượng hữu hiệu, ở các liều dùng và trong các khoảng thời gian cần thiết để đạt được kết quả mong muốn. Chẳng hạn, lượng hữu hiệu có tác dụng trị liệu của một hợp chất có thể thay đổi theo các yếu tố như tình trạng bệnh, tuổi, giới tính, và thể trọng của cá thể, và khả năng của kháng thể gây ra đáp ứng mong muốn ở cá thể. Phác đồ liều có thể được điều chỉnh để tạo ra đáp ứng trị liệu tối ưu.

Theo một phương án, hợp chất (*chẳng hạn*, virut sống đã được giảm độc lực của các phương án này) có thể được dùng theo cách thuận tiện, chẳng hạn, dưới da, trong tĩnh mạch, dùng qua đường miệng, xông hít, trong chân bì, dùng qua da, dùng trong âm đạo, dùng tại chỗ, dùng trong mũi hoặc dùng qua đường trực tràng. Tùy thuộc vào đường dùng, hoạt chất có thể được chứa trong dung dịch đệm bảo vệ (*chẳng hạn*, albumin và trehaloza, poloxamer 407/trehaloza/albumin (FTA)). Theo một phương án, chế phẩm có thể được dùng qua đường miệng. Theo một phương án khác, chế phẩm có thể được dùng trong tĩnh mạch. Theo một phương án, chế phẩm có thể được dùng trong mũi, như xông hít. Theo một phương án khác nữa, chế phẩm có thể được dùng trong chân bì bằng cách sử dụng hệ không kim tiêm (*chẳng hạn*, Pharmajet®) hoặc các hệ dùng trong chân bì khác.

Theo các phương án nhất định, chế phẩm có thể được dùng cho trẻ em hoặc thanh niên trong chất mang hoặc chất pha loãng thích hợp. Theo các phương án nhất định, vacxin virut sốt xuất huyết sống đã được giảm độc lực có thể được dùng cho trẻ em hoặc thanh niên trong hỗn hợp làm ổn định (chẳng hạn albumin và trehaloza; FTA, hoặc các hỗn hợp khác để ổn định virut sống đã được giảm độc lực. "Chất mang dược dụng" như được sử dụng ở đây cũng có thể bao gồm các chất pha loãng như nước muối và các dung dịch đệm chứa nước. Có thể cần phải kết hợp chế phẩm virut sống đã được giảm độc lực với, hoặc dùng đồng thời hợp chất này với, một nguyên liệu để ngăn ngừa sự mất hoạt tính của nó. Theo các phương án nhất định, một hoặc nhiều chế phẩm virut sốt xuất huyết sống giảm độc lực có thể được dùng cho trẻ em hoặc thanh niên theo đường dưới da hoặc trong chân bì một liều ban đầu,

theo sau là liều tăng cường. Các thể phân tán cũng có thể được điều chế trong glyxerol, polyetylen glycol lỏng, và hỗn hợp của chúng và trong dầu. Trong các điều kiện bảo quản và sử dụng thông thường, các chế phẩm này có thể chứa chất bảo quản để ngăn ngừa sự phát triển của các vi sinh vật hoặc chứa chế phẩm làm ổn định khác (*chẳng hạn*, FTA hoặc các chế phẩm làm ổn định khác dùng để làm ổn định flavivirut sống giảm độc lực hoặc thể khảm của nó).

Dược phẩm thích hợp dùng để tiêm có thể được dùng theo các phương pháp đã biết trong lĩnh vực. Chẳng hạn, dung dịch nước vô trùng (có thể hòa tan trong nước) hoặc thể phân tán và bột vô trùng để điều chế ngay lập tức dung dịch tiêm vô trùng hoặc thể phân tán có thể được sử dụng. Theo một số phương án, chế phẩm có thể là vô trùng và có thể là chất lỏng để dễ tiêm. Chất mang dược dụng có thể là dung môi hoặc môi trường phân tán chứa, chẳng hạn, nước, etanol, polyol (chẳng hạn, glyxerol, propylen glycol, và polyeteylen glycol lỏng, và các chất tương tự), và các hỗn hợp thích hợp của nó. Tính chảy thích hợp có thể được duy trì, ví dụ, bằng cách sử dụng chất phủ như lexithin, bằng cách duy trì cỡ hạt mong muốn trong trường hợp thể phân tán và bằng cách sử dụng chất hoạt động bề mặt. Việc ngăn ngừa vi sinh vật có thể đạt được bằng cách bổ sung các chất chống khuẩn hoặc chống nấm khác nhau hoặc các chất khác.

Theo một số phương án, khi bào chế, dung dịch có thể được dùng theo cách tương thích với dạng liều bào chế và với lượng sao cho có hiệu quả trị liệu. Theo các phương án này, các chế phẩm có thể được dùng một cách dễ dàng trong các dạng liều khác nhau, như các loại dung dịch tiêm được mô tả ở trên. Trong các ví dụ nhất định, virut sốt xuất huyết sống giảm độc lực và/hoặc thể khảm có thể được phân phối trong chế phẩm với tỷ lệ cụ thể, chẳng hạn 5:4:5:5, 4:4:5:5, hoặc tỷ lệ khác (*chẳng hạn*, PFU của typ huyết thanh virut sốt xuất huyết đã nêu) mà ở đó các cấu trúc virut sốt xuất huyết-3 và/hoặc virut sốt xuất huyết-4 trong chế phẩm tam giá hoặc tứ giá có pfu lớn hơn ít nhất một nửa log so với virut sốt xuất huyết-1 và/hoặc virut sốt xuất huyết-2 có mặt trong chế phẩm. Theo các phương án nhất định, virut sốt xuất huyết-3 và virut sốt xuất huyết-4 sống giảm độc lực hoặc thể khảm có nồng độ pfu lớn hơn ít nhất một log so với nồng độ pfu của virut sốt xuất huyết-1 và/hoặc virut sốt xuất

huyết-2 trong chế phẩm tứ giá. Theo một số phương án, thể khảm DEN2/4 có thể có mặt với nồng độ lớn hơn các typ huyết thanh virut sốt xuất huyết khác như virut sốt xuất huyết-1 sống giảm độc lực và/hoặc virut sốt xuất huyết-2 sống giảm độc lực.

Theo các phương án nhất định, liều duy nhất hoặc đa liều chế phẩm virut sốt xuất huyết cũng có thể được dùng cho trẻ em hoặc thanh niên khoảng 1 tuổi đến khoảng 20 tuổi. Theo một số phương án, trẻ em hoặc thanh niên có thể được điều trị bằng chế phẩm liều đơn. Theo một phương án khác, trẻ em hoặc thanh niên có thể được điều trị theo ít nhất hai liều chế phẩm virut sốt xuất huyết sống đã được giảm độc lực. Theo các phương án nhất định, trẻ em hoặc thanh niên có thể được dùng chế phẩm chứa hỗn hợp virut sốt xuất huyết-virut sốt xuất huyết ở ngày 0 và liều tăng cường trong vòng khoảng 3 tháng tính từ liều đầu tiên. Theo các phương án nhất định, trẻ em và thanh niên là các đối tượng chưa từng tiếp xúc với virut (âm tính huyết thanh) chưa từng được tiếp xúc với virut sốt xuất huyết. Theo một phương án khác, trẻ em và thanh niên có thể được tiếp xúc trước đó với virut sốt xuất huyết và/hoặc sự nhiễm virut sốt xuất huyết (dương tính huyết thanh). Theo các phương án này, trẻ em âm tính huyết thanh và/hoặc thanh niên âm tính huyết thanh có thể được điều trị ở ngày 0 và sau đó nhận liều tăng cường trong vòng 6 tháng, trong vòng 5 tháng, trong vòng 4 tháng, trong vòng 3 tháng hoặc ít hơn tính từ liều đầu tiên để tạo ra đáp ứng miễn dịch gia tăng đối với các virut sốt xuất huyết. Theo các phương án nhất định, trẻ em có thể là trẻ em từ khoảng 2 tuổi đến khoảng 17 tuổi. Theo một phương án khác, trẻ em là từ 2 tuổi đến 17 tuổi.

Theo một phương án khác, dạng dung dịch, dạng xịt, sol khí hoặc dạng xông hít dùng cho mũi có thể được sử dụng để phân phối chế phẩm virut sốt xuất huyết sống đã được giảm độc lực cho đối tượng. Các chế phẩm nhất định có thể bao gồm tá dược, chẳng hạn, manitol, lactoza, tinh bột, magie stearat, natri sacarin, xenluloza, magie cacbonat loại dùng cho dược phẩm và các chất tương tự.

Dược phẩm có thể được bào chế với chất mang mà bảo vệ hoạt chất khỏi bị bài tiết nhanh chóng từ cơ thể, như các chế phẩm hoặc lớp bao giải phóng theo thời gian. Các chất mang như vậy có thể bao gồm các chế phẩm giải phóng có kiểm soát, như, nhưng không giới hạn ở, hệ phân phối được bao vi nang, và các polyme tương

thích sinh học dễ phân hủy sinh học, như etylen vinyl axetat, polyanhydrit, axit polyglycolic, polyorthoeste, axit polylactic và các chất mang khác là đã biết.

Theo các phương án nhất định, khoảng liều của thể khảm virut sốt xuất huyết-virut sốt xuất huyết sống giảm độc lực, virut sốt xuất huyết sống giảm độc lực hoặc thể khảm flavivirut có thể từ khoảng 10^2 đến khoảng 10^6 PFU được dùng ban đầu và tùy ý, sau đó bởi ít nhất lần dùng thứ hai trong vòng 30 ngày hoặc lên tới 12 tháng sau đó, nếu cần. Theo các phương án nhất định, đối tượng có thể nhận liều dùng kép gồm chế phẩm đơn giá, nhị giá, tam giá hoặc tứ giá được bộc lộ ở đây trong đó chế phẩm này là hỗn hợp một thành phần duy nhất hoặc có thành phần định trước gồm các typ huyết thanh virut sốt xuất huyết khác nhau.

Hiển nhiên là, đối với trẻ em hoặc thanh niên cụ thể bất kỳ hoặc đối với trẻ em có độ tuổi bất kỳ, phác đồ liều cụ thể có thể được điều chỉnh theo nhu cầu của mỗi cá nhân. Theo một số phương án, phác đồ liều được điều chỉnh để thích ứng với các đối tượng vị thành niên, chẳng hạn như trẻ em khoảng 20 tuổi hoặc trẻ hơn, hoặc trẻ em khoảng 18 tuổi hoặc trẻ hơn hoặc khoảng 15 tuổi hoặc trẻ hơn hoặc khoảng 12 tuổi hoặc trẻ hơn hoặc khoảng 9 tuổi hoặc trẻ hơn hoặc khoảng 6 tuổi hoặc trẻ hơn, hoặc khoảng 3 tuổi hoặc trẻ hơn hoặc khoảng 1,5 tuổi hoặc trẻ hơn. Theo các phương án nhất định, trẻ em có thể là 2 tuổi đến 17 tuổi. Theo một phương án khác, trẻ em có thể là 4 đến 16 tuổi.

Theo các phương án nhất định, chế phẩm sinh miễn dịch được bộc lộ ở đây có thể được dùng cho trẻ em hoặc thanh niên, ở dạng một hoặc nhiều liều. Theo một số phương án, chế phẩm sinh miễn dịch được bộc lộ ở đây có thể được dùng cho đối tượng ở dạng liều đơn, hoặc ở dạng hai liều hoặc nhiều hơn trong khoảng thời gian định trước, bao gồm nhưng không giới hạn ở, trong khoảng 6 tháng, trong khoảng 120 ngày, trong khoảng 90 ngày, trong khoảng 80 ngày, trong khoảng 70 ngày, trong khoảng 60 ngày, trong khoảng 50 ngày, trong khoảng 40 ngày, trong khoảng 30 ngày, trong khoảng 20 ngày, trong khoảng 10 ngày, trong khoảng 5 ngày, hoặc ít hơn hoặc vào cùng ngày trong vòng vài giờ hoặc vài phút hoặc được dùng đồng thời ở các vị trí giải phẫu giống hoặc khác nhau. Theo một số phương án, các chế phẩm sinh miễn dịch được bộc lộ ở đây có thể được dùng trong vòng khoảng 90 ngày mỗi loại, trong

vòng khoảng 60 ngày mỗi loại, trong vòng khoảng 30 ngày mỗi loại, và trong vòng ít hơn khoảng 30 ngày mỗi loại. Theo một số phương án, chế phẩm được bộc lộ ở đây có thể được dùng dưới da hoặc trong chân bì cho đối tượng. Việc dùng ở hai hoặc nhiều hơn hai vị trí giải phẫu có thể bao gồm tổ hợp việc dùng bất kỳ bao gồm theo cùng phương thức ở hai hoặc nhiều hơn hai vị trí giải phẫu hoặc theo hai phương thức khác nhau mà bao gồm hai vị trí giải phẫu riêng biệt. Theo các phương án này, hai hoặc nhiều hơn hai vị trí giải phẫu có thể bao gồm các chi khác nhau hoặc các vùng khác nhau của cơ thể. Theo các phương án nhất định, hai liều chế phẩm vacxin có thể được đưa liên tiếp vào đối tượng ở cùng hoặc nhiều vị trí giải phẫu để, chẳng hạn, bảo vệ chống lại tất cả các typ huyết thanh virut sốt xuất huyết (*chẳng hạn*, bảo vệ chéo) ở ngày 0. Theo các phương án nhất định, chế phẩm sinh miễn dịch được bộc lộ ở đây có thể bao gồm, nhưng không giới hạn ở, liều đơn mà gây ra phản ứng sinh miễn dịch đối với cả bốn typ huyết thanh (*chẳng hạn*, chế phẩm tứ giá) được dùng cho đối tượng, có khả năng gây ra đáp ứng miễn dịch đối với tất cả các typ huyết thanh virut sốt xuất huyết. Theo các phương án nhất định, đối tượng có thể nhận chế phẩm liều đơn ở đó liều đơn này có khả năng gây ra đáp ứng miễn dịch thích đáng đối với cả bốn typ huyết thanh virut sốt xuất huyết. Theo một phương án khác, chế phẩm sinh miễn dịch có thể bao gồm các typ huyết thanh virut sốt xuất huyết sống đã được giảm độc lực kết hợp với các tác nhân sinh miễn dịch khác chống lại các flavivirut khác (*chẳng hạn*, virut Zika, virut gây viêm não Nhật Bản, virut West Nile, virut gây viêm não St Louis, virut gây sốt vàng hoặc các virut khác). Theo một phương án nhất định, vacxin chống lại virut sốt xuất huyết được bộc lộ ở đây có thể được sử dụng để làm giảm sự nhiễm virut có liên quan khác, chẳng hạn, nhiễm virut Zika.

Theo một số phương án, dược phẩm được bộc lộ ở đây có thể được sử dụng để làm tăng đáp ứng miễn dịch của chế phẩm mục tiêu ở trẻ em hoặc thanh niên bằng cách phối hợp chế phẩm sinh miễn dịch chống virut sốt xuất huyết với một tác nhân để tăng cường đáp ứng tế bào $CD8^+$ T hoặc đáp ứng miễn dịch khác chống lại một hoặc nhiều typ huyết thanh virut sốt xuất huyết (chẳng hạn một hoặc nhiều thành phần cấu trúc hoặc không cấu trúc của virut sốt xuất huyết) ở trẻ em hoặc thanh niên

nhận dược phẩm chống lại virut sốt xuất huyết này.

Việc dùng tại chỗ được thực hiện thông qua kem, gel, dung dịch rửa, v.v. được dùng tại chỗ chứa các chất ức chế serin proteaza với lượng hữu hiệu có tác dụng trị liệu. Việc dùng qua da được thực hiện bằng cách dùng kem, dung dịch rửa, gel, v.v. có khả năng cho phép các chất ức chế serin proteaza thấm qua da và đi vào dòng máu. Ngoài ra, bơm thẩm thấu có thể được sử dụng để dùng chế phẩm. Liều cần thiết sẽ thay đổi theo tình trạng bệnh cụ thể cần điều trị, phương pháp dùng và tốc độ thanh thải của phân tử khỏi cơ thể.

Các phương án khác liên quan đến phương pháp làm giảm khả năng mất hoạt tính của virut sống đã được giảm độc lực bao gồm, nhưng không giới hạn ở, kết hợp một hoặc nhiều virut sống giảm độc lực với chế phẩm có khả năng làm giảm sự mất hoạt tính của virut sống đã được giảm độc lực (chẳng hạn flavivirut). Các chế phẩm này có thể bao gồm, nhưng không giới hạn ở, một hoặc nhiều tác nhân protein; một hoặc nhiều sacarit hoặc tác nhân polyol; và tùy ý, một hoặc nhiều copolyme khối EO-PO, trong đó chế phẩm này có thể làm giảm sự mất hoạt tính hoặc làm ổn định virut sống giảm độc lực.

Theo các phương án nhất định, chế phẩm được dự tính ở đây có thể được loại nước hoặc hydrat hóa một phần hoặc hoàn toàn. Theo một phương án khác, tác nhân protein làm ổn định được dự tính để dùng trong dược phẩm hoặc chế phẩm không phải dược phẩm ở đây có thể bao gồm, nhưng không giới hạn ở, lactalbumin, albumin huyết thanh người, albumin huyết thanh người tái tổ hợp (ecombinant người albumin huyết thanh- rHSA), albumin huyết thanh bò (bovine albumin huyết thanh - BSA), các albumin huyết thanh khác hoặc các thành viên thuộc họ gen albumin. Các tác nhân sacarit hoặc polyol có thể bao gồm, nhưng không giới hạn ở, monosacarit, disacarit, các rượu đường, trehaloza, sucroza, maltoza, isomaltoza, cellibioza, gentiobioza, laminariboza, xylobioza, mannobioza, lactoza, fructoza, sorbitol, mannitol, lactitol, xylitol, erythritol, rafinoza, amylza, xyclodextrin, chitosan, hoặc xenluloza. Theo các phương án nhất định, chất hoạt động bề mặt có thể bao gồm, nhưng không giới hạn ở, chất hoạt động bề mặt không ion chẳng hạn như alkyl poly(etylen oxit), copolyme của poly(etylen oxit) và poly(propylen oxit) (copolyme

khối EO-PO), poly(vinyl pyroloidon), alkyl polyglucosit (chẳng hạn như sucroza monostearat, lauryl diglucosit, hoặc sorbitan monolaureat, octyl glucosit và dexyl maltosit), các rượu béo (rượu xetyl hoặc rượu oleyl), hoặc cocamit (cocamit MEA, cocamit DEA và cocamit TEA).

Theo một phương án khác, chất hoạt động bề mặt có thể bao gồm, nhưng không giới hạn ở, poloxamer 407 (chẳng hạn Pluronic F127®), poloxamer 335, 338 hoặc 238 thay cho hoặc ngoài poloxamer 407, hoặc các copolyme khối EO-PO khác có đặc tính tương tự như F127®.

Theo một số phương án, chế phẩm vacxin có thể bao gồm, nhưng không giới hạn ở, một hoặc nhiều tác nhân protein mà là albumin huyết thanh; một hoặc nhiều tác nhân sacarit mà là trehaloza; và một hoặc nhiều tác nhân polyme hoạt động bề mặt chẳng hạn như copolyme khối EO-PO, poloxamer 407 hoặc cụ thể hơn, Pluronic F127®.

Theo một phương án khác, các chế phẩm để làm ổn định các virut sống có thể bao gồm một hoặc nhiều Flavivirut sống, một hoặc nhiều tác nhân hydrat cacbon, và một hoặc nhiều axit amin hoặc các dẫn xuất muối, este hoặc amit của nó. Theo một phương án khác, chế phẩm để dùng trong bản mô tả này làm ổn định Flavivirut sống giảm độc lực để dùng trong thương mại. Theo một số phương án, chế phẩm này còn bao gồm chất đệm. Theo các phương án này, chất đệm có thể bao gồm nhưng không giới hạn ở, dung dịch muối đệm phosphat (phosphate buffered saline - PBS). Theo các phương án này, chất đệm có thể bao gồm ít nhất một natri clorua (NaCl), mononatri và/hoặc dinatri phosphat (Na_2HPO_4), kali clorua (KCl), và kali phosphat (KH_2PO_4). Theo một số phương án, chất đệm của chế phẩm có thể bao gồm natri clorua ở nồng độ 25 mM đến 200 mM. Theo một phương án khác, chế phẩm được bộc lộ ở đây có thể bao gồm ure và/hoặc các tác nhân thích hợp khác MSG.

Theo một số phương án, Flavivirut sống giảm độc lực chẳng hạn như virut sốt xuất huyết có thể được làm ổn định trong chế phẩm mà bao gồm nhưng không giới hạn ở, HSA tái tổ hợp có nồng độ từ 0,1% đến 0,2% (w/v); và/hoặc sucroza có nồng độ từ khoảng 4,0% đến khoảng 6,0% (w/v); và/hoặc mannitol có nồng độ từ khoảng 2% đến 4% (w/v); và/hoặc alanin có nồng độ từ khoảng 8,0 mM đến khoảng 22,0

mM; và/hoặc methionin có nồng độ từ khoảng 1,0 mM đến khoảng 5,0 mM; và/hoặc MSG có nồng độ từ khoảng 8,0 mM đến 12,0 mM; và/hoặc ure có nồng độ từ khoảng 0,1% đến khoảng 0,3% (w/v). Theo các phương án nhất định, chế phẩm có thể bao gồm HSA tái tổ hợp, trehaloza, mannitol, alanin, methionin, MSG và ure. Theo một phương án khác, chế phẩm làm ổn định có thể bao gồm HSA ở nồng độ từ khoảng 0,1% đến khoảng 0,2% (w/v); nồng độ trehaloza từ khoảng 4% đến khoảng 6% (w/v); nồng độ mannitol từ khoảng 2% đến khoảng 4% (w/v); trong đó nồng độ alanin là từ 8 mM đến 22 mM; trong đó nồng độ methionin là từ 1 mM đến 5 mM; trong đó nồng độ MSG là từ 8 mM đến 12 mM; và trong đó nồng độ ure là từ 0,1% đến 0,3% (w/v). Các chế phẩm nhất định để làm ổn định virut sống giảm độc lực có thể bao gồm, nhưng không giới hạn ở, HSA tái tổ hợp, sucroza, alanin và ure. Theo các phương án này, nồng độ HSA có thể là từ khoảng 0,1% đến khoảng 0,2% (w/v); nồng độ sucroza có thể là từ khoảng 4% đến khoảng 6% (w/v); nồng độ alanin có thể là từ khoảng 8,0 mM đến khoảng 22 mM; và nồng độ ure có thể là từ khoảng 0,1% đến khoảng 0,3% (w/v). Các chế phẩm làm ổn định khác có thể bao gồm HSA tái tổ hợp, sucroza, methionin và ure. Nồng độ HSA tái tổ hợp có thể là từ khoảng 0,1% đến 0,2% (w/v); nồng độ sucroza có thể là từ khoảng 4,0 % đến khoảng 6,0 % (w/v); nồng độ methionin có thể là từ khoảng 1,0 mM đến khoảng 5,0 mM; và nồng độ ure có thể là từ khoảng 0,1% đến khoảng 0,3% (w/v). Theo một phương án khác, chế phẩm làm ổn định có thể bao gồm HSA tái tổ hợp, sucroza, arginin và ure trong đó nồng độ HSA tái tổ hợp có thể là từ 0,1% đến 0,2% (w/v); nồng độ sucroza có thể là từ 4% đến 6% (w/v); nồng độ arginin có thể là từ 10 mM đến 50 mM; và nồng độ ure có thể là từ 0,1% đến 0,3% (w/v). Các chế phẩm làm ổn định khác có thể bao gồm HSA tái tổ hợp, trehaloza, arginin và ure trong đó nồng độ HSA tái tổ hợp là từ khoảng 0,1% đến 0,2% (w/v); nồng độ trehaloza là từ khoảng 4% đến 6% (w/v); nồng độ arginin là từ khoảng 10 mM đến 50 mM; và nồng độ ure là từ khoảng 0,1% đến 0,3% (w/v). Theo một phương án khác, chế phẩm làm ổn định có thể bao gồm HSA tái tổ hợp, trehaloza, MSG và ure. Theo các phương án này, nồng độ HSA tái tổ hợp có thể là từ khoảng 0,1% đến khoảng 0,2% (w/v); nồng độ trehaloza có thể là từ khoảng 4,0% đến khoảng 6,0 % (w/v); trong đó nồng độ MSG là từ 8 mM đến 12 mM; và trong đó nồng độ ure là từ 0,1% đến 0,3% (w/v).

Một số phương án ở đây liên quan đến chế phẩm virut sống giảm độc lực được loại nước một phần hoặc hoàn toàn vì lý do vận chuyển hoặc các lý do khác. Theo các phương án này, chế phẩm có thể được loại nước 20 % hoặc nhiều hơn; 30% hoặc nhiều hơn; 40% hoặc nhiều hơn; 50% hoặc nhiều hơn; 60% hoặc nhiều hơn; 70 % hoặc nhiều hơn; 80% hoặc nhiều hơn; hoặc 90% hoặc nhiều hơn. Theo các phương án này, chế phẩm vacxin virut có thể được loại nước và hydrat hóa lại trong chế phẩm làm ổn định bất kỳ đã biết trước khi sử dụng chế phẩm dược dụng cho trẻ em hoặc thanh niên.

Theo các phương án nhất định, đối tượng có thể là động vật có vú chẳng hạn như người hoặc động vật thú y và/hoặc động vật đã được thuần hóa hoặc gia súc hoặc động vật hoang dã. Theo các phương án nhất định, chế phẩm sinh miễn dịch theo sáng chế có thể là hữu hiệu để gây miễn dịch các đối tượng vị thành niên, chẳng hạn như người trẻ em mà 20 tuổi hoặc trẻ hơn. Mặc dù tình trạng kỹ thuật liên quan đến các vacxin virut sốt xuất huyết/chế phẩm sinh miễn dịch hiện nay (chẳng hạn thể khảm virut sốt xuất huyết/virut gây sốt vàng) được sử dụng ở trẻ em báo cáo hiệu lực và/hoặc tính sinh miễn dịch kém đối với trẻ em 9 tuổi hoặc trẻ hơn, chế phẩm sinh miễn dịch được bộc lộ ở đây có thể gây ra đáp ứng miễn dịch hữu hiệu ở trẻ em từ khoảng 1 tuổi đến khoảng 17 tuổi hoặc khoảng 1 tuổi đến khoảng 9 tuổi, hoặc lớn hơn. Chế phẩm sinh miễn dịch được bộc lộ ở đây thể hiện hiệu lực và tính sinh miễn dịch lớn hơn so với các chế phẩm đã biết.

Theo một phương pháp được nêu làm ví dụ, chế phẩm sinh miễn dịch được bộc lộ ở đây có mặt cả bốn typ huyết thanh virut sốt xuất huyết (*chẳng hạn,* chế phẩm tứ giá), có thể được dùng cho đối tượng, ở đó chế phẩm sinh miễn dịch có thể gây ra đáp ứng miễn dịch đối với cả bốn typ huyết thanh virut sốt xuất huyết ở đối tượng. Theo các phương án nhất định, chế phẩm sinh miễn dịch theo sáng chế có thể bao gồm tổ hợp gồm các polynucleotit và/hoặc polypeptit của các cấu trúc typ huyết thanh virut sốt xuất huyết khác nhau hoặc virut sốt xuất huyết sống giảm độc lực được bộc lộ ở đây. Theo một phương án khác, chế phẩm sinh miễn dịch có thể bao gồm polynucleotit có trình tự axit nucleic mã hóa typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực được biểu thị bởi SEQ ID NO: 9 (Pre-

Master), 11 (MVS), 13 (WVS) hoặc 15 (BVS); polynucleotit khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2 có trình tự axit nucleic mã hóa protein không cấu trúc từ typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực và protein cấu trúc từ typ huyết thanh virut sốt xuất huyết-1, được biểu thị bởi SEQ ID NO: 1 (Pre-Master), 3 (MVS), 5 (WVS) hoặc 7 (BVS); polynucleotit khảm virut sốt xuất huyết-3/virut sốt xuất huyết-2 có trình tự axit nucleic mã hóa protein không cấu trúc từ typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực và protein cấu trúc từ typ huyết thanh virut sốt xuất huyết-3, được biểu thị bởi SEQ ID NO: 17 (Pre-Master), 19 (MVS), 21 (WVS) hoặc 23 (BVS); và polynucleotit khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 có trình tự axit nucleic mã hóa protein không cấu trúc từ typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực và protein cấu trúc từ typ huyết thanh virut sốt xuất huyết-4 được biểu thị bởi SEQ ID NO: 25 (Pre-Master), 27 (MVS), 29 (WVS) hoặc 31 (BVS). Các polynucleotit này có thể chứa hoặc không chứa các polypeptit khác nhau mà chúng mã hóa. Theo các phương án nhất định, chế phẩm sinh miễn dịch được bộc lộ ở đây có thể bao gồm một hoặc nhiều polypeptit tương ứng với typ huyết thanh virut sốt xuất huyết 2 sống được cải biến giảm độc lực được biểu thị bởi SEQ ID NO: 10 (Pre-Master), 12 (MVS), 14 (WVS) hoặc 16 (BVS); một hoặc nhiều polypeptit tương ứng với polynucleotit khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2 được biểu thị bởi SEQ ID NO: 2 (Pre-Master), 4 (MVS), 6 (WVS) hoặc 8 (BVS); một hoặc nhiều polypeptit tương ứng với polynucleotit khảm virut sốt xuất huyết-3/virut sốt xuất huyết-2 được biểu thị bởi SEQ ID NO: 18 (Pre-Master), 20 (MVS), 22 (WVS) hoặc 24 (BVS); và/hoặc một hoặc nhiều polypeptit tương ứng với polynucleotit khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 được biểu thị bởi SEQ ID NO: 26 (Pre-Master), 28 (MVS), 30 (WVS) hoặc 32 (BVS).

Nồng độ của các polynucleotit và các polypeptit khác nhau được bộc lộ ở đây (*chẳng hạn*, TDV-1, TDV-2, TDV-3 và TDV-4) có thể là như nhau, hoặc các typ huyết thanh virut sốt xuất huyết nhất định có thể có mặt trong chế phẩm sinh miễn dịch nhiều hơn các typ huyết thanh khác, tùy thuộc vào nhu cầu hoặc sự bùng phát của một hoặc nhiều typ huyết thanh. Theo các phương án này, các nồng độ hoặc tỷ lệ

khác nhau có thể được biểu thị dưới dạng log nồng độ PFU so với nhau, như được minh họa trên các Fig. 11A-11B. Nồng độ có thể được áp dụng cho các chế phẩm tứ giá khác nhau của chế phẩm sinh miễn dịch được bộc lộ ở đây dựa trên tỷ lệ của TDV-1 đối với TDV-2 đối với TDV-3 đối với TDV-4, chẳng hạn. Nồng độ có thể bao gồm, nhưng không giới hạn ở log nồng độ PFU là, 3:3:3:3, 4:4:4:4, 5:5:5:5, 5:1:5:5, 3:2:3:3, 4:2:4:4, 5:3:5:5, 4:3:4:4, 5:4:5:5, 4:4:5:5, hoặc tỷ lệ nồng độ bất kỳ của typ huyết thanh virut sốt xuất huyết bất kỳ, tùy thuộc vào, chẳng hạn, số lượng typ huyết thanh có mặt trong chế phẩm, đáp ứng định trước và hiệu lực mong muốn, mà sẽ được người có hiểu biết trung bình trong lĩnh vực sẵn sàng thừa nhận dựa trên bản mô tả này.

Theo các phương án nhất định, nồng độ của virut sốt xuất huyết-2 sống giảm độc lực có thể thấp hơn 0,5 1, 1,5 2, 2,5 3 hoặc 4 log PFU so với một hoặc nhiều typ huyết thanh virut sốt xuất huyết khác trong chế phẩm sinh miễn dịch để gây ra đáp ứng miễn dịch cân bằng hơn đối với cả bốn typ huyết thanh virut sốt xuất huyết và làm giảm sự can thiệp của virut. Mặc dù tình trạng kỹ thuật liên quan đến vacxin virut sốt xuất huyết/chế phẩm sinh miễn dịch hiện thời báo cáo hiệu lực và/hoặc tính sinh miễn dịch thấp hơn đối với virut sốt xuất huyết-2, chế phẩm sinh miễn dịch được bộc lộ ở đây đề xuất dạng bào chế mà có thể tạo ra đáp ứng miễn dịch hữu hiệu đối với cả bốn typ huyết thanh virut sốt xuất huyết ở nhiều hơn 60% đối tượng và trong các trường hợp nhất định ở khoảng 85% (xem, *chẳng hạn*, các Fig. 15-16), bao gồm virut sốt xuất huyết-2, mặc dù là virut sốt xuất huyết-2 sống giảm độc lực có thể thấp hơn ít nhất một log PFU so với các typ huyết thanh virut sốt xuất huyết khác. Chế phẩm sinh miễn dịch theo sáng chế thể hiện hiệu lực và tính sinh miễn dịch tốt hơn so với các chế phẩm đã biết đối với đối tượng 20 tuổi hoặc trẻ hơn và tạo ra đáp ứng miễn dịch cân bằng hơn đối với cả bốn typ huyết thanh virut sốt xuất huyết ở lứa tuổi bất kỳ. Theo các phương án nhất định, tỷ lệ chuyển dạng huyết thanh có thể được đánh giá để phân tích hiệu lực của chế phẩm nhất định theo thời gian sau khi dùng một hoặc nhiều liều vacxin sống giảm độc lực.

Theo các phương án nhất định, chế phẩm sinh miễn dịch chống lại virut sốt xuất huyết có thể bao gồm một hoặc nhiều thể khảm virut sốt xuất huyết-1/virut sốt

xuất huyết-2 có nồng độ từ khoảng 1,0 x 10^3 đến khoảng 5 x 10^5 PFU; virut sốt xuất huyết-2 sống giảm độc lực có nồng độ từ khoảng 1,0 x 10^3 đến khoảng 5 x 10^5 PFU; thể khảm virut sốt xuất huyết-3/virut sốt xuất huyết-2 có nồng độ từ khoảng 5,0 x 10^3 đến khoảng 5 x 10^5 PFU; và/hoặc thể khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 có nồng độ từ khoảng 1,0 x 10^4 đến khoảng 5 x 10^6 PFU. Theo một số phương án, chế phẩm tứ giá có thể bao gồm thể khảm virut sốt xuất huyết-3/virut sốt xuất huyết-2 với nồng độ lớn hơn ít nhất một nửa đến một log so với thể khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2 và virut sốt xuất huyết-2 sống giảm độc lực ở chế phẩm tam giá hoặc tứ giá trong khi virut sốt xuất huyết-4/virut sốt xuất huyết-2 có thể có nồng độ lớn hơn ít nhất một nửa đến một log so với thể khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2 và virut sốt xuất huyết-2 sống giảm độc lực ở chế phẩm tam giá hoặc tứ giá. Theo các phương án nhất định, nồng độ của virut sốt xuất huyết-3/virut sốt xuất huyết-2 và thể khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 có thể là giống nhau hoặc nồng độ thể khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 có thể lớn hơn virut sốt xuất huyết-3/virut sốt xuất huyết-4 ở chế phẩm nhị giá, tam giá hoặc tứ giá được dự tính ở đây. Theo một số phương án đối tượng được điều trị bởi chế phẩm và phương pháp được bộc lộ ở đây có thể được điều trị 2 lần trong một năm, hằng năm, 18 tháng hoặc phác đồ tương tự tùy thuộc vào địa điểm của đối tượng và kế hoạch di chuyển chẳng hạn. Theo các phương án nhất định, chế phẩm sinh miễn dịch chống lại virut sốt xuất huyết có thể bao gồm một hoặc nhiều thể khảm virut sốt xuất huyết-1/virut sốt xuất huyết-2 có nồng độ khoảng 2,0 x 10^4 PFU; virut sốt xuất huyết-2 sống giảm độc lực có nồng độ khoảng 5,0 x 10^3 PFU; thể khảm virut sốt xuất huyết-3/virut sốt xuất huyết-2 có nồng độ khoảng 1 x 10^5 PFU; và/hoặc thể khảm virut sốt xuất huyết-4/virut sốt xuất huyết-2 có nồng độ khoảng 3.0 x 10^5 PFU.

Theo các phương án nhất định, trẻ em hoặc thanh niên nhận chế phẩm được bộc lộ ở đây có thể được đánh giá qua một khoảng thời gian về sự phát triển triệu chứng hoặc dấu hiệu bất kỳ có liên quan đến nhiễm virut sốt xuất huyết hoặc nhiễm flavivirut khác. Chẳng hạn, trẻ em hoặc thanh niên có thể được đánh giá sự khởi phát sốt xuất huyết hoặc triệu chứng hoặc dấu hiệu khác có liên quan đến nhiễm virut sốt

xuất huyết.

Các phương pháp trị liệu

Theo một phương án của sáng chế, phương pháp đề xuất để gây ra đáp ứng miễn dịch đối với typ huyết thanh virut sốt xuất huyết bằng cách sử dụng chế phẩm đơn giá, nhị giá, tam giá hoặc tứ giá chứa các cấu trúc virut sống giảm độc lực và/hoặc khảm được dự tính ở đây.

Các phương án theo sáng chế được minh họa kỹ hơn bởi các ví dụ không giới hạn sau đây, mà nó không được hiểu là nhằm giới hạn theo cách bất kỳ dựa trên phạm vi của nó. Ngược lại, cần phải hiểu rõ rằng phải viện đến các phương án khác, các cải biến, và các phương án tương đương của nó mà, sau khi đọc bản mô tả ở đây, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này có thể tự đề xuất mà không đi trệch khỏi tinh thần của sáng chế hoặc phạm vi của yêu cầu bảo hộ kèm theo.

Ví dụ thực hiện sáng chế

Các ví dụ sau đây được bao hàm để biểu thị các phương án nhất định được đưa ra trong sáng chế. Người có hiểu biết trung bình trong lĩnh vực sẽ hiểu rằng, các kỹ thuật được bộc lộ trong phần ví dụ tiếp theo dưới đây đại diện cho các kỹ thuật được phát hiện để thực hiện tốt trong thực tiễn được bộc lộ ở đây, và vì vậy có thể được xem như tạo thành các phương thức được ưu tiên để thực hiện sáng chế. Tuy nhiên, với sự bộc lộ này, người có hiểu biết trung bình trong lĩnh vực sẽ hiểu rằng, nhiều thay đổi có thể được thực hiện trong các phương án cụ thể mà chúng được bộc lộ và vẫn thu được kết quả tương tự hoặc giống hệt mà không đi trệch ra khỏi tinh thần và phạm vi được nêu ở đây.

Ví dụ 1

Theo một số phương pháp được nêu làm ví dụ, chế phẩm được sử dụng để tạo ra hạt virut mà được đề cập đến ở đây như là "hạt virut chủng gốc (MVS)" được bộc lộ. Các chế phẩm này có thể có nguồn gốc từ một hoặc nhiều virut sốt xuất huyết sống giảm độc lực, chẳng hạn như DEN-1, DEN-2, DEN-3, và DEN-4. Theo các phương pháp nhất định, chế phẩm có thể có nguồn gốc từ một hoặc nhiều virut sốt xuất huyết sống giảm độc lực mà bao gồm nhưng không giới hạn ở, các cấu trúc cụ

thể được bộc lộ ở đây được đề cập đến như là TDV-1, TDV-2, TDV-3, và TDV-4 (chẳng hạn trước đây được gọi là DENVax). Theo một phương pháp được nêu làm ví dụ khác, các chiến lược được dùng để tạo ra và mô tả đặc trưng các chế phẩm này được đề xuất. Theo một phương án khác nữa, chế phẩm virut sốt xuất huyết tứ giá và đặc tính di truyền và kiểu hình của các chế phẩm này được đề xuất. Các cấu trúc này có thể được sử dụng để tạo ra chế phẩm sinh miễn dịch dùng để điều trị cho trẻ em hoặc thanh niên từ 1 tuổi đến 20 tuổi.

Sản xuất và phân tích các virut TDV tiền chủng gốc

Các thủ tục nhất định được thực hiện để tạo ra hạt virut sốt xuất huyết tiền chủng gốc, chẳng hạn như khuếch đại và tinh chế theo chuỗi đối với virut sốt xuất huyết (*chẳng hạn*, TDV). Trước tiên, virut TDV được tạo dẫn xuất lại bằng cách chuyển nhiễm ARN virut được phiên mã từ ADN bổ trợ của TDV tái tổ hợp với chiều dài đầy đủ vào tế bào được chứng nhận dùng trong sản xuất (*chẳng hạn*, tế bào Vero), tạo ra hạt virut P1 (cấy chuyển 1). Sau đó, bốn virut P1 từ mỗi virut sốt xuất huyết-1 đến virut sốt xuất huyết-4 được khuếch đại và vết tan được tinh chế để thu được hạt P7 vacxin tiền chủng gốc (xem Bảng 1). Các thử nghiệm nhất định được thực hiện để phân tích các giống cấy chuyển của virut sốt xuất huyết. Chẳng hạn, việc phân tích trình tự bộ gen với chiều dài đầy đủ đã chứng tỏ rằng cả bốn hạt virut P2 (cấy chuyển 2) đều giống hệt về mặt di truyền với virut vacxin ứng viên loại dùng trong nghiên cứu, có nguồn gốc từ loại dùng cho nghiên cứu, virut nguyên bản tương đồng của chúng. Kiểu hình vết tan ban đầu cũng được giữ lại ở virut P2. Sáu virut đã được tinh chế vết (P3 A-F) được phân lập đối với mỗi typ huyết thanh của virut sốt xuất huyết (*e.g.*, TDV1-4) từ các hạt P2, và mỗi vết tan đã phân lập được tinh chế vết tan trực tiếp thêm hai lần nữa. Sản phẩm tinh tinh chế vết tan lần thứ ba (P5) của mỗi virut được khuếch đại hai lần (P6 A-F và P7 A-F) ở các tế bào Vero để tạo ra các hạt TDV P7 tiền chủng gốc tiềm năng (Bảng 1).

Bảng 1 Ví dụ về tái tạo dẫn xuất cGMP của Virut TDV ở các tế bào WCB-Vero

Cấy chuyển	Sản xuất/tinh chế hạt	Đặc tính
P1	Chuyển nhiễm WCB-Vero với ARN virut đã được phiên mã	Chuẩn độ vết
P2	Khuếch đại virut P1	Trình tự bộ gen đầy đủ
P3	Chọn 6 vết tan (A-F)/typ huyết thanh từ thử nghiệm vết tan P2	Tinh chế vết tan
P4	Chọn vết tan A-F từ thử nghiệm P3	Tinh chế vết tan
P5	Chọn vết tan A-F từ thử nghiệm P4	Tinh chế vết tan
P6	Khuếch đại vết tan P5 A-F	Chuẩn độ vết
P7	Hạt tiền chủng gốc: Khuếch đại P6 A-F	Trình tự bộ gen đầy đủ, TaqMAMA, kiểu hình vết tan
P8*	MVS: Khuếch đại hạt virut P7 được chọn lựa	Mô tả đầy đủ đặc tính di truyền và kiểu hình
P9	WVS: Khuếch đại hạt virut chủng gốc P8	Trình tự bộ gen đầy đủ, TaqMAMA
P10	BVS: Khuếch đại hạt virut giống sản xuất P9	Trình tự bộ gen đầy đủ, TaqMAMA
* Một hạt P7 tối ưu (A, B, C, D, E, hoặc F) được lựa chọn dựa trên phân tích di truyền và vết tan để tạo ra P8 MVS		

Một số thử nghiệm được thực hiện tiếp để mô tả đặc trưng hạt P7 TDV, chẳng hạn như phân tích trình tự bộ gen và kiểu hình vết tan của hạt P7, và so sánh với hạt P2 (Bảng 2). Kiểu hình vết tan của virut P7 thường tương tự với kiểu hình của hạt P2. Trong một số thí nghiệm được nêu làm ví dụ, độ chuẩn virut được giám sát. Độ chuẩn virut đạt trên 6,0 log pfu/ml đối với hầu hết các hạt P7, ngoại trừ đối với 5 virut. Đọc trình tự gen của nhiều hơn 60 hạt virut vacxin ứng viên sau 10 hoặc nhiều hơn mười chuỗi cấy chuyển ở tế bào Vero không nhận dạng được trường hợp đột

biến ngược nào ở NS1-53 và NS3-250 của ba yếu tố quyết định giảm độc lực chính của vectơ di truyền DENV-2 PDK-53, gợi ý rằng 2 locus này khá bền trong các hạt virut vacxin ứng viên. Thử nghiệm khuếch đại không ghép đôi nhạy cảm (TaqMAMA) được phát triển để xác định một cách chính xác tỷ lệ đột biến ngược ở locus 5'NCR-57 bởi RT-PCR thời gian thực. Trong một số nghiên cứu, tỷ lệ đột biến ngược 5'NCR-57 của toàn bộ 24 hạt P7 được xác định bởi TaqMAMA. Tùy thuộc vào nồng độ của ARN virut đưa vào đối với mỗi virut trong thử nghiệm, giới hạn độ nhạy cảm của TaqMAMA nằm trong khoảng từ 0,01% đến 0,07% đột biến ngược, mà nó nhạy cảm hơn nhiều so với giới hạn độ nhạy cảm đột biến ngược 10-30% có thể phát hiện được bởi phương pháp phân tích trình tự bộ gen liên ứng. Dữ liệu thu được cho thấy 15 trong số 24 virut P7 có tỷ lệ đột biến ngược tối thiểu hoặc không thể phát hiện được (< 0,07%), một virut (TDV-3-D) có tỷ lệ đột biến ngược gần như 100%, và 8 virut (chẳng hạn TDV-1, 1 TDV-2, 2 TDV-3, và 4 TDV-4) có tỷ lệ đột biến ngược một phần nằm trong khoảng từ 0,08% đến 12,85% (Bảng 2). Việc phân tích trình tự bộ gen với chiều dài đầy đủ được thực hiện đối với 16 trong số 24 virut P7 với mức độ đột biến ngược của 5'NCR57 thấp như được xác định bởi TaqMAMA. Tất cả các virut được đọc trình tự vẫn có hai yếu tố xác định độc lực giảm TDV (NS1-53, NS3-250), và đều có các đột biến bổ sung thu được mà không có mặt trong các dòng ADN bổ trợ tái tổ hợp được xây dựng ban đầu (Bảng 2). Trong một chế phẩm vacxin mục tiêu được nêu làm ví dụ, TDV-1-A, TDV-2-F, TDV-3-F, và TDV-4-F được lựa chọn làm hạt tiền chủng gốc mục tiêu đối với mỗi typ huyết thanh do kiểu gen và kiểu hình vết tan của chúng gần giống nhất so với kiểu gen và kiểu hình vết tan của các dạng tái tổ hợp vacxin được thiết kế ban đầu. TDV-1-A, TDV-2-F, và TDV-4-F có hai đột biến không cùng nghĩa, và TDV-3-F có một đột biến không cùng nghĩa. Bằng chứng này gợi ý rằng các đột biến bổ sung được quan sát ở 4 hạt tiền chủng gốc này không gây ra mối bận tâm về độ an toàn hoặc các biến đổi về tính sinh miễn dịch đối với các virut này. Các hạt tiền chủng gốc này được khuếch đại tiếp để tạo ra MVS (hạt chủng gốc, có tên gọi là P8, Bảng 1).

Các phương pháp được nêu làm ví dụ được đề cập đến ở đây sử dụng ARN

virut được phiên mã in-vitro được tinh chế từ plasmit ADN bổ trợ được tạo dòng làm nguồn tinh khiết để chuyển nhiễm tế bào Vero đã được chứng thực dùng trong vacxin để tạo ra virut dùng trong vacxin. Việc tinh chế vết tan theo chuỗi và phân tích trình tự bộ gen đầy đủ được kết hợp vào quy trình sản xuất để đảm bảo hạt vacxin được sản xuất ra có độ tinh khiết tối ưu và độ ổn định về mặt di truyền. Sáu virut đã tạo dòng được bào chế ở dạng các hạt tiền chủng gốc tiềm năng cho mỗi typ huyết thanh của TDV. Qua phân tích bộ gen, bao gồm đọc trình tự TaqMAMA và bộ gen đầy đủ, cũng như mô tả đặc trưng của kiểu hình vết tan ở virut, các hạt tiền chủng gốc được lựa chọn để cải tiến sự sản xuất chủng virut giống gốc đối với mỗi typ huyết thanh (các typ huyết thanh 1-4). Các hạt tiền chủng gốc được lựa chọn có tỷ lệ đột biến ngược không thể phát hiện được (<0,01% hoặc <0,07%) ở locus 5'NCR-57, với 1 hoặc 2 phần tử thế axit amin trong bộ gen của chúng, và vẫn có kiểu hình vết tan nhỏ như được quan sát trước đó.

Bảng 2. Đặc trưng của hạt tiền chủng gốc (P7)

Virut	Dòng[a]	TaqMAMA[b]	Log_{10} pfu/ml	Vết tan[c]	Các đột biến được nhận biết trong bộ gen[d]
TDV-1	**A**	**	6,85	P2	NS2A-116 I-L, NS2B-92 E-D, một đột biến câm
	B	*	6,93	P2	nd[e]
	C	*	6,93	D	nd
	D	**	7,02	D	C-67 K-A; một đột biến câm
	E	0,57%	7,28	P2	nd
	F	**	7,18	P2	E473 T-M; một đột biến câm
TDV-2	A	0,03%	6,33	P2	NS1-341 K-N
	B	*	6,33	P2	E-305 K-T, hai đột biến câm
	C	*	5,84	L	NS4A-18 T-A, bốn đột biến câm
	D	0,08%	6,20	P2	NS2B-99 I-L, một 3'NCR

	E	0,03%	6,31	P2	prM-52 K-E, NS5-412 I-V, hai đột biến câm
	F	**	6,15	P2	prM-52 K-E, NS5-412 I-V
TDV-3	A	*	6,00	P2	NS5-200 K-N, một đột biến câm, một 3'NCR
	B	0,05%	6,27	P2	NS2A-33 I-T, NS2A-59 M-T
	C	0,30%	6,25	P2	nd
	D	100,00%	6,27	P2	nd
	E	0,31%	6,00	P2	nd
	F	**	6,30	P2	E-223 T-S, một đột biến câm
TDV-4	A	0,47%	5,60	P2	E323 K-R/K, NS2B-21 L-F/L, NS2B-39 T-S, một đột biến câm
	B	*	5,65	D	NS2A-126 A-V; NS4A-5 N-D;NS5-383 K-R, một đột biến câm
	C	4,50%	5,90	P2	nd
	D	12,85%	5,97	D	nd
	E	0,52%	6,85	S	prM-85 E-D, NS2B-45 T-A, NS5-320 M-T, NS5-551 E-G, hai đột biến câm
	F	0,02%	6,93	S	NS2A-66 D-G, NS4A-21 A-V, bốn đột biến câm

[a] Các virut đã tạo dòng (bằng cách tinh chế vết tan theo chuỗi) được lựa chọn để phát triển tiếp MVS được in đậm.

[b] *: Tỷ lệ đột biến ngược < 0,07% (giới hạn phát hiện). **: Tỷ lệ đột biến ngược < 0,01% (giới hạn phát hiện)

[c] kiểu hình vết tan: P2: tương tự như virut P2; L = lớn hơn virut P2, D = kích thước tương tự, nhưng dường như hơi khác về độ sắc nét của vết tan; S = nhỏ hơn P2.

[d]Các phần tử thế khác với dòng ADN bổ trợ TDV được thiết kế. Đột biến axit amin được liệt kê với phần còn lại của protein virut và là những thay đổi (đột biến kiểu dại). Tổng số đột biến câm ở gen cấu trúc và gen không cấu trúc của mỗi hạt được liệt kê. Các đột biến ở vùng không mã hóa (NCR) cũng được ghi lại.

[e] nd = Không được thực hiện. Các dòng này có tỷ lệ đột biến ngược 5'NCR-57 (bởi TaqMAMA) cao hơn các dòng khác, vì vậy được không được phân tích trình tự tiếp theo.

Ví dụ 2

Theo một phương pháp khác được nêu làm ví dụ, chế phẩm chứa chủng virut giống gốc, chủng virut giống sản xuất và chủng virut với số lượng lớn cũng như các đặc tính di truyền và kiểu hình của chúng được mô tả. Các chế phẩm này được đề xuất để sản xuất các nguyên liệu lâm sàng và các nguồn cung cấp vacxin thương mại cơ bản. Bước tinh chế vết theo chuỗi và phân tích trình tự bộ gen đầy đủ được kết hợp vào quy trình sản xuất để đảm bảo chế phẩm chứa hạt vacxin có độ an toàn và độ ổn định di truyền tối ưu để sản xuất nguyên liệu thử nghiệm lâm sàng.

Sản xuất và quản lý chất lượng trong quá trình sản xuất MVS, WVS, và BVS

Ở một số nghiên cứu, MVS của 4 TDV được sản xuất bằng cách khuếch đại hạt P7 tiền chủng gốc trong tế bào Vero đã được chứng nhận. Trong các nghiên cứu khác, MVS được dùng để tạo ra một lượng lớn WVS ở các tế bào. Ngoài ra, các chủng gốc BVS của TDV được khuếch đại từ WVS và được sử dụng cho các thử nghiệm lâm sàng ở người. Quản lý chất lượng để giải phóng sản phẩm được thực hiện theo một số phương pháp được nêu làm ví dụ, bao gồm bước thử nghiệm tất cả các MVS, WVS, và BVS đối với tính đồng nhất, độ chuẩn lây nhiễm, khả năng vô trùng, mycoplasma, và các tác nhân ngẫu nhiên in vitro và in vivo. Tất cả các hạt được cho qua thử nghiệm tính đồng nhất virut bằng cách sử dụng thử nghiệm RT-PCR đặc hiệu typ huyết thanh, mà nó thể hiện sự khuếch đại dương tính tương ứng với typ huyết thanh của nó và âm tính đối với typ huyết thanh khác loại (dữ liệu không được chỉ ra). Không phát hiện thấy mycoplasma hoặc các tác nhân ngẫu nhiên được phát hiện ở các chủng gốc MVS, WVS, BVS.

Theo một phương pháp được nêu làm ví dụ, chế phẩm sinh miễn dịch được

bộc lộ ở đây có mặt cả bốn typ huyết thanh virut sốt xuất huyết (*chẳng hạn*, tứ giá), có thể được dùng cho trẻ em hoặc thanh niên, trong đó chế phẩm sinh miễn dịch có thể gây ra đáp ứng miễn dịch chống lại cả bốn typ huyết thanh virut sốt xuất huyết ở trẻ em hoặc thanh niên.

Phân tích gen của MVS, WVS, và BVS

Theo các phương pháp được nêu làm ví dụ nhất định, sau khi tạo ra MVS từ tiền MVS (P7), các chủng đã được lựa chọn ở trên được tạo ra và ARN virut tương ứng này được đọc trình tự lại. Việc phân tích trình tự bộ gen với chiều dài đầy đủ cho thấy MVS đối với TDV-1 là giống hệt với hạt tiền chủng gốc của nó, trong khi WVS và BVS sau đó có 2 phần tử thế nữa ở E-483 và NS4B-108 (xem các bảng 2 và 3). Việc thế Ala ở E-483 thể hiện một phần của kiểu gen ở MVS, nhưng trở thành kiểu gen trội ở BVS. TDV-2 và TDV-3 là giống hệt với các hạt tiền chủng gốc tương ứng của chúng (Bảng 2 và 3). TDV-2 MVS giống hệt với hạt tiền chủng gốc của nó, và WVS và BVS có 2 các đột biến bổ sung ở NS4A-36 và NS4B-111. Cả hai đột biến này đều không hoàn chỉnh ở WVS và là kiểu gen chính ở BVS. MVS của TDV-3 lại giống hệt với hạt tiền chủng gốc, nhưng WVS và BVS chứa một thay thế axit amin khác nữa ở NS4A-23. TDV-4 MVS có được đột biến axit amin khác nữa, ở locus NS2A-99 (kiểu gen kết hợp từ Lys đến Lys/Arg) trong quá trình sản xuất MVS (bảng 3). WVS và BVS của nó vẫn giữ được kiểu gen kết hợp NS2A-99 Lys/Arg, và BVS có kiểu gen kết hợp thêm NS4B-238 Ser/Phe. Kết quả trình tự liên ứng cũng khẳng định rằng MVS, WVS cũng như BV vẫn giữ được ba yếu tố quyết định di truyền của tính độc giảm ở các locus 5'NCR-57, NS1-53, và NS3-250. Phân tích locus tính độc giảm ít ổn định nhất bởi TaqMAMA chứng tỏ rằng tỷ lệ đột biến ngược 5'NCR-57 nằm trong khoảng từ <0,7% đến 0,13% trong số MVS, ≤0,07% trong số WVS, và nằm trong khoảng từ <0,07 đến 0,21% trong số BVS. Tỷ lệ đột biến ngược 3% ở locus 5'NCR-57 được xem là tỷ lệ tối đa cho phép để chấp nhận một lô vacxin (Bảng 3).

Bảng 3. Thế nucleotit và axit amin ở hạt TDV

TDV	**Nucleotit**	**Axit amin**	**Tiền chủng gốc**	**MVS[a]**	**WVS[a]**	**BVS[a]**
TDV-1	2384 G-C	E-483 Gly-Ala	-	-	**Gly/Ala**	**Ala**
	3823 A-C	NS2A-116 Ile-Leu	Leu	Leu	Leu	Leu
	4407 A-T	NS2B-92 Glu-Asp	Asp	Asp	Asp	Asp
	7148 C-T	NS4B-108 Thr-Ile	-	-	**Ile**	**Ile**
	7311 A-G	Đột biến câm	G	G	G	G
	TaqMAMA 5'NCR-57 đột biến ngược %[b]		--	-	-	-
TDV-2	592 A-G	prM-52 Lys-Glu	Glu	Glu	Glu	Glu
	6481 G-C	NS4A-36 Ala-Pro	-	-	**Ala/Pro**	**Pro**
	7156 C-T	NS4B-111 Leu-Phe	-	-	**Leu/Phe**	**Phe**
	8803 A-G	NS5-412 Ile-Val	Val	Val	Val	Val
	TaqMAMA 5'NCR-57 đột biến ngược %[b]		--	-	0,07%	0,21%
TDV-3	1603 A-T	E-223 Thr-Ser	Ser	Ser	Ser	Ser
	6436 G-A	NS4A-23 Asp-Asn	-	-	**Asn**	**Asn**
	7620 A-G	Đột biến câm	G	G	G	G
	TaqMAMA 5'NCR-57 đột biến ngược %[b]		--	-	-	-
TDV-4	225 A-T	Đột biến câm	T	T	T	T
	3674 A-G	NS2A-66 Asp-Gly	Gly	Gly	Gly	Gly
	3773 A-A/G	NS2A-99 Lys-Lys/Arg	-	**Lys/Arg**	**Lys/Arg**	**Lys/Arg**

	5391 C-T	Đột biến câm	T	T	T	T
	6437 C-T	NS4A-21 Ala-Val	Val	Val	Val	Val
	7026 T-C	Đột biến câm	T/C	T/C	T/C	T/C
	7538 C-C/T	NS4B-238 Ser-Ser/Phe	-	-	**Ser/Phe**	**Ser/Phe**
	9750 A-C	Đột biến câm	C	C	C	C
	TaqMAMA 5'NCR-57 đột biến ngược %[b]		-	0,13%	-	-
[a] In đậm: Thay đổi bắt đầu ở chủng gốc MVS.						
[b] "--" biểu thị tỷ lệ đột biến ngược < 0,01% (giới hạn phát hiện), "-" biểu thị tỷ lệ đột biến ngược <0,07% (giới hạn phát hiện)						

Phân tích trình tự bộ gen đầy đủ bộc lộ rằng đột biến axit amin khác nữa được biểu lộ ở TDV-4 MVS, trong khi ba lô TDV MVS khác vẫn có trình tự bộ gen liên ứng của các hạt tiền chủng gốc của chúng. Từ hạt P1 thành hạt tiền chủng gốc (P7), chỉ 1 hoặc 2 đột biến không cùng nghĩa xuất hiện trong hạt đã nêu. Từ hạt P1 thành hạt MVS (P8), 2 đến 7 phần tử thế nucleotit giống hệt trong hạt TDV bất kỳ đã nêu và chỉ 2 đến 3 trong số các phần tử thế này tạo ra sự thay đổi axit amin. Không đột biến câm nào trong MVS là thuộc 5' hoặc 3'NCR mà có thể ảnh hưởng đến sự sao chép virut. Chỉ có sự thay đổi ở prM-52 Lys-Glu của TDV-2, và sự thay thế ở NS2A-66 Asp-Gly của TDV-4 là những thay đổi có tính bảo toàn. Đột biến NS2A-66 của TDV-4 là ở phần khung không cấu trúc của DENV-2 PDK-53. Mặc dù locus NS2A-66 thường là Asp trong số các chủng DENV-2 khác nhau, ở DENV-4 nó thường là Gly. Có thể sự thay đổi từ Asp thành Gly ở TDV-4 là để phù hợp với TDV-4 ở các tế bào Vero. Các gốc đột biến TDV-2 prM-52 ở phần đầu tận cùng C của prM mà nó được tách loại từ hạt virut trưởng thành. Ở một số phương pháp được nêu làm ví dụ, mô tả đặc trưng về kiểu hình được thực hiện để xác nhận rằng không đột biến nào trong các hạt MVS làm thay đổi đáng kể kiểu hình giảm độc lực của vacxin.

Virut TDV chứng tỏ độ ổn định di truyền cao trong suốt quá trình sản xuất. Ba locus giảm độc lực DENV-2 PDK-53 xác định nằm trong 5'NCR, NS1-53, và NS3-250 vẫn ổn định trong trình tự gen liên ứng trong khi cấy chuyển theo chuỗi TDV từ tiền chủng gốc thành các chế phẩm vacxin với lượn lớn. TaqMAMA có độ nhạy cao của locus 5'NCR-57 thể hiện đột biến ngược tối thiểu hoặc không thể phát hiện được ở MVS, WVS (P9/giống sản xuất), và BVS (hạt virut chủ yếu với số lượng lớn dùng cho vacxin) của typ huyết thanh virut sốt xuất huyết. Tỷ lệ đột biến ngược 5'NCR-57 của các chế phẩm TDV BVS (tương đương P10) thấp hơn đáng kể so với tỷ lệ đột biến ngược 5'NCR-57 mà được suy ra trong các ứng viên vacxin loại dùng cho nghiên cứu sau 10 lần cấy chuyển trong các tế bào Vero (tỷ lệ đột biến ngược 4-74%). Chiến lược để sản xuất các hạt TDV ở quy mô lớn được đề xuất ở đây tạo ra hạt vacxin ổn định về mặt di truyền mà nó vẫn có các gen đánh dấu giảm độc lực trong các virut vacxin ứng viên.

Tất cả các cấu trúc virut TDV có thể được bảo quản trong chất đệm làm ổn định, chẳng hạn, FTA, poloxamer 407 F127® (0,01% đến khoảng 3,0% w/v) và khoảng 5% đến khoảng 50% (w/v) trehaloza, và khoảng 0,01% đến khoảng 3,0% albumin (chẳng hạn rHSA)

Độ an toàn và tính sinh miễn dịch *in vivo*

Trong ví dụ này, đã chứng minh rằng các chế phẩm được nêu làm ví dụ là an toàn sau khi tiêm dưới da và về cơ bản trơ về mặt miễn dịch học. Bốn chế phẩm được nêu làm ví dụ khác được lựa chọn để thử nghiệm ở chuột nhắt là như sau (dữ liệu không được chỉ ra).

Chế phẩm 1: 15% trehaloza, 2% F-127, 1% rHSA

Chế phẩm 2: 15% trehaloza, 2% F-127, 1% rHSA, 1mM $CaCl_2$/ 0,5mM $MgSO_4$

Chế phẩm 3: 15% trehaloza, 2% F-127, 1% rHSA, 0,5% chitosan

Chế phẩm 4: 22,5% trehaloza, 3% F-127, 1,5% rHSA

Chế phẩm 5: PBS

Kiểu hình vết tan của TDV MVS

Theo một phương pháp được nêu làm ví dụ, kiểu hình vết tan của TDV MVS được so sánh với virut sốt xuất huyết kiểu dại và virut thể khảm tương đồng loại dùng cho nghiên cứu của chúng ở các tế bào Vero (Fig. 2). Tất cả các MVS của TDV-1, -2, và -3 tạo ra vết tan nhỏ hơn đáng kể so với các thể tương đồng kiểu dại của chúng và rất giống (chênh lệch trong khoảng 0,4-mm) với các virut tương đồng loại nghiên cứu của chúng ở các tế bào Vero. TDV-4 MVS cũng nhỏ hơn đáng kể so với DENV-4 kiểu dại, nhưng lớn hơn một chút (chênh lệch 0,9 mm) so với thể khảm D2/4-V bắt nguồn từ chủng gốc trong phòng thí nghiệm.

Fig. 2 là biểu đồ được nêu làm ví dụ minh họa kích thước vết tan của TDV MVS trái ngược với virut kiểu dại đối chứng và các virut ứng viên dùng làm vacxin loại nghiên cứu. Đường kính vết tan trung bình (mm) ± SD (khoảng sai số) của vết tan virut ở các tế bào Vero trong điều kiện phủ agaroza được đo ở ngày 9 sau gây nhiễm. Virut DEN kiểu dại, được biểu thị bởi cột màu đen, và các virut ứng viên dùng làm vacxin loại nghiên cứu đã được công bố trước đó, được biểu thị bởi cột màu trắng, được bao gồm để đối chứng và so sánh với hạt vacxin chủng gốc TDV được biểu thị bởi cột màu xám.

Độ nhạy với nhiệt độ của TDV MVS

Theo một phương pháp khác được nêu làm ví dụ, độ nhạy với nhiệt độ được thử nghiệm ở các tế bào Vero đối với TDV MVS và được so với kiểu dại tương đồng của chúng và virut vacxin khảm loại nghiên cứu ban đầu. DENV-3 16562 kiểu dại (WT) không nhạy cảm với nhiệt độ. Typ huyết thanh virut sốt xuất huyết kiểu dại 1 và typ huyết thanh virut sốt xuất huyết-4 nhạy cảm vừa phải với nhiệt độ ở 39°C (độ chuẩn ở 39°C thấp hơn khoảng 1,0 $\log_{10}$ pfu/ml so với ở 37°C, Fig. 3). Typ huyết thanh virut sốt xuất huyết kiểu dại-2 16681 nhạy cảm nhất với nhiệt độ trong số các virut sốt xuất huyết kiểu dại được thử nghiệm, và tạo ra độ chuẩn giảm 100 lần ở 39°C. TDV-1, -2, và -3 nhạy với nhiệt độ bằng với virut vacxin khảm tương đồng loại nghiên cứu ban đầu của chúng (Fig. 2). Độ chuẩn ở 39°C giảm trong khoảng từ 2,0 đến 3,0 $\log_{10}$ pfu/ml đối với các chủng TDV này. TDV-4 cũng nhạy với nhiệt độ, thể hiện sự giảm độ chuẩn 5 lần. Tuy nhiên, D2/4-V loại nghiên cứu ban đầu thể hiện

độ chuẩn giảm khoảng 10 lần. TDV-4 MVS ổn định cuối cùng chứa poloxamer 407, F127® (và các tác nhân khác để ổn định các chế phẩm này (FTA hoặc chế phẩm nhị giá, albumin và trehaloza)), đã được chứng minh trước đó là gia tăng độ ổn định nhiệt của virut sốt xuất huyết. Sự có mặt của F127® trong TDV-4 MVS dường như góp phần vào độ nhạy với nhiệt độ thấp rõ rệt của virut trong thử nghiệm nuôi cấy Vero. Trong một thí nghiệm riêng rẽ, độ nhạy với nhiệt độ của chủng TDV-4 bắt nguồn từ MVS với sự vắng mặt của F127® cũng được đánh giá. Để loại bỏ poloxamer 407, F127® từ chủng này, ARN virut được phân lập từ chế phẩm virut chứa chủ yếu TDV-4 và được chuyển nhiễm vào tế bào Vero. Virut TDV-4 này tỏ ra là nhạy với nhiệt độ giống như chủng nghiên cứu D2/4 V (độ chuẩn giảm 1,5 $\log_{10}$ pfu/ml) ở ngày thứ 3 sau gây nhiễm với sự vắng mặt của poloxamer 407, F127® (Fig. 3).

Fig. 3 minh họa biểu đồ được nêu làm ví dụ minh họa độ nhạy với nhiệt độ của TDV MVS. Virut sốt xuất huyết kiểu dại và các virut ứng viên dùng làm vacxin loại nghiên cứu đã được công bố trước đó được dùng để so sánh. TDV-4 MVS chứa poloxamer 407, F127® khác nữa mà có thể che giấu kết quả độ nhạy với nhiệt độ của virut trong thử nghiệm này. Một thử nghiệm riêng rẽ phân tích một TDV-4 đại diện với sự vắng mặt của F127® cũng được bao hàm. Độ chuẩn trung bình ± SD (khoảng sai số) của các virut được sao chép ở các tế bào Vero ở 37°C hoặc 39°C.

Sự sao chép TDV MVS ở các tế bào C6/36 ở muỗi

Theo một số phương pháp được nêu làm ví dụ, TDV MVS được phát triển ở các tế bào C6/36 để xác nhận sự duy trì kiểu hình giảm độc lực *in vitro* của chúng, biết rằng các virut vacxin khảm loại nghiên cứu vẫn có kiểu hình giảm độc lực của bộ khung virut DENV-2 PDK53 ở các tế bào muỗi này. So với virut sốt xuất huyết kiểu dại, TDV-1, TDV-2 và TDV-4 MVS thể hiện sự giảm sinh trưởng đáng kể (giảm ít nhất 3 $\log_{10}$ pfu/ml) ở các tế bào C6/36 ở ngày 6 sau gây nhiễm (post infection - pi) (Fig. 4). TDV-3 MSV cũng thể hiện sự sinh trưởng giảm so với DENV-3 16562 kiểu dại, nhưng sự giảm này là không đáng kể như vậy (giảm 1-2 $\log_{10}$ pfu/ml). Tuy nhiên, độ chuẩn C6/36 của các lô hạt TDV-3 tương tự (độ chênh lệch trong khoảng 1 $\log_{10}$ pfu/ml) với độ chuẩn C6/36 của virut vacxin D2/3-V

khamr loại nghiên cứu ban đầu.

Fig. 4 minh họa biểu đồ được nêu làm ví dụ lập biểu đồ sự sinh trưởng hạn chế của TDV MVS (cột màu xám) ở các tế bào C6/36 so với virut sốt xuất huyết kiểu dại (cột màu đen) và virut vacxin loại nghiên cứu (cột màu trắng). Độ chuẩn trung bình ± SD (khoảng sai số) của virut đã được sao chép trong các tế bào C6/36 6 ngày sau gây nhiễm.

Tỷ lệ gây nhiễm, phát tán, và truyền virut ở muỗi nguyên vẹn

Theo một số phương pháp được nêu làm ví dụ, tỷ lệ gây nhiễm và phát tán TDV được so sánh với virut sốt xuất huyết kiểu dại bố mẹ của chúng. Trong các thí nghiệm được nêu làm ví dụ nhất định, các thí nghiệm gây nhiễm theo đường miệng được thực hiện ở muỗi *Ae. aegypti*. Bột máu đã gây nhiễm được chuẩn độ ngược để xác định độ chuẩn virut và chỉ có những thí nghiệm với độ chuẩn virut tương tự trong bột máu (chênh lệch nhỏ hơn 1 $\log_{10}$ pfu/ml) giữa virut sốt xuất huyết bố mẹ và TDV đối với mỗi typ huyết thanh được sử dụng để so sánh trong bảng 4. TDV-1, TDV-2, và D2 PDK-53-VV45R loại nghiên cứu không lây nhiễm cho muỗi nhờ ăn qua đường miệng, mà nó khác biệt đáng kể ($p < 0,0001$) với các virut bố mẹ của chúng, DENV-1 16007 (gây nhiễm 44%) và DENV-2 16681 (gây nhiễm 43,3%). Do không có muỗi nào bị gây nhiễm bởi TDV-1 và -2, có từ rất ít đến không có sự phát tán liên quan đến hai virut vacxin này. Trong khi TDV-4 gây nhiễm một số muỗi nhờ ăn qua đường miệng (2 trong số 55), tỷ lệ gây nhiễm thấp hơn đáng kể ($p < 0,05$) so với virut kiểu dại bố mẹ của chúng, DENV-4 1036 (8 trong số 50). TDV-3 không gây nhiễm cho bất kỳ con muỗi nào ở hai thí nghiệm với độ chuẩn virut trong bột máu là 5,2±0,02 $\log_{10}$ pfu/ml (Table 4), và trong một thí nghiệm riêng rẽ với độ chuẩn virut trong bột máu là 6,0 $\log_{10}$ pfu/ml, chỉ có 1 trong số 30 con muỗi là bị gây nhiễm (dữ liệu không được chỉ ra). Tuy nhiên, virut sốt xuất huyết-3 16562 kiểu dại cũng có tỷ lệ gây nhiễm rất thấp (8%) ở 5,2 $\log_{10}$ pfu/ml, và tỷ lệ này không gia tăng trong một thí nghiệm riêng rẽ với độ chuẩn virut trong bột máu cao hơn ở 6,2 $\log_{10}$ pfu/ml (3%, 1 con dương tính trong số 30 con muỗi, dữ liệu không được chỉ ra). Mặc dù virut sốt xuất huyết-3 và virut sốt xuất huyết-4 kiểu dại (WT) có tỷ lệ gây nhiễm thấp hơn đáng kể so với virut sốt xuất huyết-1 và virut sốt xuất huyết-2 kiểu dại, độ

chuẩn virut trung bình ở muỗi đã gây nhiễm là tương tự (3,1 đến 3,9 log_{10} pfu/con muỗi). Ngược lại, độ chuẩn TDV-4 từ hai con muỗi đã gây nhiễm đều đạt tối thiểu (0,7 log_{10} pfu/con muỗi), nhỏ hơn 1.000 lần so với độ chuẩn từ các con muỗi đã gây nhiễm bởi typ huyết thanh virut sốt xuất huyết-4 1036 kiểu dại (3,9 ± 1,5 pfu/con muỗi).

Đối với các con muỗi mà đã bị gây nhiễm, sự phát tán ra khỏi ruột giữa có thể được đánh giá bằng cách xác định xem liệu virut có mặt trong chân hay không. Bốn virut bố mẹ DENV tạo ra tỷ lệ phát tán nằm trong khoảng từ 36,3% đến 62,5%, và độ chuẩn virut trung bình của chúng (tính bằng log_{10} pfu) từ các chân nằm trong khoảng 0,9±0,3 đến 2,2±0,7 (ngoại trừ các mẫu âm tính). Không con muỗi nào trong số hai con TDV-4 đã được gây nhiễm tạo ra sự phát tán virut đến chân (bảng 4). Mặc dù virut đã phát tán được phát hiện ở chân, không virut nào trong số bốn virut sốt xuất huyết kiểu dại được phát hiện ở nước bọt của muỗi được gây nhiễm qua đường miệng, gợi ý rằng các điều kiện cho ăn qua đường miệng có thể không được đủ nhạy cảm để xác định tỷ lệ truyền của các DENV này. Vì vậy, theo các phương pháp khác được nêu làm ví dụ, sự gây nhiễm muỗi nhân tạo cực kỳ nghiêm ngặt bởi sự chủng ngừa IT trực tiếp được thực hiện sau đó (bảng 4). Ngoại trừ TDV-4, tất cả các virut (kiểu dại và TDV) đạt được tỷ lệ gây nhiễm 100% *Ae. aegypti* đã chủng ngừa IT. Dịch chủng ngừa TDV-4 có độ chuẩn virut thấp hơn đáng kể so với ba dịch chủng ngừa virut khác, nhưng nó vẫn được gây nhiễm đáng kể cho 70% muỗi đã chủng ngừa. Mặc dù tỷ lệ gây nhiễm toàn thân cao đạt được bằng cách chủng ngừa IT, cả bốn virut TDV đều có tỷ lệ truyền thấp hơn đáng kể ($p < 0,005$) hoặc không phát hiện được (0-10%) so với virut sốt xuất huyết kiểu dại (43-87%, bảng 4). Virut TDV thể hiện sự gây nhiễm và phát tán rất ít hoặc không gây nhiễm và phát tán sau khi ăn qua đường miệng, và các kết quả IT cực kỳ nghiêm ngặt đã xác nhận khả năng truyền tối thiểu của các virut TDV này ở *Ae. aegypti*.

Bảng 4: Tỷ lệ gây nhiễm, phát tán, và truyền virut ở muỗi nguyên vẹn

Virut	Ăn qua đường miệng					Chủng ngừa IT				
	Bột máu[a]	Gây nhiễm[b]	Độ chuẩn toàn thân[c]		Sự phát tán[e]	Dịch chủng ngừa	Gây nhiễm[b]	Độ chuẩn toàn thân[c]	Nước bọt[f]	
	Giá trị trung bình±SD	% (P/N)	Giá trị trung bình±SD	p^d	% (P/N)[f]	pfu/liều	% (P/N)	Giá trị trung bình±SD	% (P/N)	p^d
DENV-1 16007	6,6	44,0% (11/25)	3,6 ± 1,5		36,3% (4/11)	53,9	100% (30/30)	4,7 ± 0,48	43% (13/30)	
TDV-1	6,9	0% (0/30)	Không khả dụng	<0,0001	Không khả dụng	67,8	100% (30/30)	3,4 ± 0,39	10% (3/30)	<0,005
DENV-2 16681	6,6	43,3% (13/30)	3,1 ± 1,5		38,5% (5/13)	67,8	100% (30/30)	5,2 ± 0,34	87% (26/30)	
D2 PDK53-VV45R	6,4	0% (0/30)	Không khả dụng	<0,0001	Không khả dụng	56,4	100% (30/30)	4,0 ± 0,20	0% (0/30)	<0,0001
TDV-2	6,4	0% (0/30)	Không khả dụng	<0,0001	Không khả dụng	52,7	100% (30/30)	3,5 ± 0,27	7% (2/30)	<0,0001
DENV-3 16562	5,2	8% (2/25)	3,8 ± 0,2		50% (1/23)	34,0	100% (30/30)	4,2 ± 0,50	67% (20/30)	
TDV-3	5,2 ± 0,02	0% (0/50)	Không khả dụng	0,108	Không khả dụng	37,3	100% (30/30)	3,3 ± 0,36	3% (1/30)	<0,0001
DENV-4 1036	5,8 ± 0,5	16% (8/50)	3,9 ± 1,5		62,5% (5/8)	69,4	100% (30/30)	5,2 ± 0,45	70% (21/30)	
TDV-4	5,4 ± 0,4	3,6% (2/55)	0,7 ± 0,0	0,033	0% (0/2)	11,8	70% (21/30)	1,1 ± 0,46	0% (0/21)	<0,0001

[a] Độ chuẩn virut hoặc giá trị trung bình±độ lệch chuẩn nếu từ nhiều hơn 1 thí nghiệm trong bột máu (log_{10} pfu/ml) bằng cách chuẩn độ ngược

[b] Tỷ lệ virut được phát hiện trong các cơ thể muỗi. *P/N* = số muỗi dương tính/tổng số muỗi

[c] độ chuẩn virut trung bình ± độ lệch chuẩn (log_{10} pfu/muỗi) trong cơ thể muỗi, chủ các mẫu dương tính là được sử dụng để tính

[d] Phân tích thống kê chênh lệch giữa DENV kiểu dại và TDV bằng xác suất Fisher Exact

[e] Tỷ lệ virut được phát hiện ở chân của muỗi bị gây nhiễm dương tính

[f] Tỷ lệ virut được phát hiện ở nước bọt của muỗi bị gây nhiễm dương tính. Được sử dụng để xác định hiệu quả truyền

Năng lực của vectơ là thành phần an toàn quan trọng đối với virut vacxin flavivirut sống giảm độc lực. Trước đây, virut DENV-2 PDK-53-VV45R loại nghiên

cứu và dẫn xuất đột biến ngược kiểu dại được thử nghiệm trong *Ae. aegypti,* và nhận thấy rằng đột biến giảm độc lực NS1-53-Asp là yếu tố quyết định tính trội đối với sự sao chép giảm ở muỗi. Hai locus giảm độc lực chính của vacxin DENV-2 PDK-53, nucleotit 5'NCR-57-T và NS3-250-Val, cũng thể hiện một số tác dụng ức chế đối với sự sao chép ở muỗi, vì vậy tạo ra các hạn chế dư thừa khác đối với năng lực vectơ của muỗi. Một số phương pháp được nêu làm ví dụ được mô tả ở đây được dùng để kiểm tra sự gây nhiễm IT và qua đường miệng ở muỗi và sự sao chép đối với cả bốn chủng TDV. TDV-1, -2, và -3 không gây nhiễm bất kỳ muỗi *Ae. aegypti* nào bằng cách gây nhiễm qua đường miệng (bảng 4). TDV-4 gây nhiễm chỉ 3,6% muỗi được tiếp xúc qua đường miệng, mức độ gây nhiễm thấp hơn đáng kể so với mức độ gây nhiễm của DENV-4 kiểu dại với độ chuẩn trung bình sao chép ở cơ thể muỗi thấp hơn độ chuẩn trung bình sao chép ở muỗi đã gây nhiễm DENV-4 kiểu dại. Bất ngờ là, TDV-4 đã được phát hiện ở chân của muỗi đã gây nhiễm, cho thấy TDV-4 không có khả năng phát tán từ ruột giữa của muỗi sau khi gây nhiễm qua đường miệng. Tỷ lệ gây nhiễm đối với TDV-1, -2, và -4 đều nhỏ hơn đáng kể so với virut kiểu dại tương ứng của chúng, nhưng mức chênh lệch này là không đáng kể giữa TDV-3 và DENV-3 16562 kiểu dại do tỷ lệ gây nhiễm rất thấp đối với cả hai virut này. So với các chủng DENV kiểu dại khác được đánh giá trong *Ae. aegypti* cùng được thu nhận từ tỉnh Mae Sot, Thái Lan, các chủng virut sốt xuất huyết kiểu dại bố mẹ được dùng để tạo ra TDV dường như có tỷ lệ gây nhiễm và phát tán thấp hơn bằng cách gây nhiễm qua đường miệng. DENV-1 PUO359, DENV-2 PUO218, DENV-3 PaH881/88, và DENV-4 1288 kiểu dại được dùng để tạo ra các vacxin ChimeriVax-DEN trên cở sở vacxin virut gây sốt vàng (YF) 17D có tỷ lệ gây nhiễm nằm trong khoảng 47-77%. Ngược lại, vacxin YF 17D không thể gây nhiễm *Ae. aegypti*. Mặc dù các chủng ChimeriVax chứa prM-E từ DENV kiểu dại gây nhiễm ở tỷ lệ cao này, ChimeriVax nãy vẫn có kiểu hình giảm độc lực ở muỗi của bộ khung sao chép YF 17D của chúng. Các kết quả được nêu ở đây cũng chỉ ra rằng sự giảm độc lực đối với muỗi của bộ khung DENV-2 PDK-53 được duy trì ở các chủng TDV. Ngoài ra, bằng cách sử dụng các chủng virut sốt xuất huyết kiểu dại với khả năng gây nhiễm thấp hơn ở muỗi của các cấu trúc chứa trong chế phẩm được mô tả ở đây tạo ra một đặc điểm an toàn khác.

Theo một phương pháp được nêu làm ví dụ, kết quả gây nhiễm qua đường miệng minh họa rằng TDV có khả năng gây nhiễm và khả năng phát tán tối thiểu bởi muỗi. Ngoài ra, các thí nghiệm gây nhiễm IT nhạy cảm và nghiêm ngặt hơn đã được thực hiện để phân tích kỹ hơn tiềm năng của TDV để được truyền nhiễm bởi *Ae. aegypti*. Kết quả IT đã chứng tỏ rằng cả bốn virut TDV có khả năng truyền nhiễm bởi muỗi ở mức không thể phát hiện dược hoặc tối thiểu so với các virut tương đương kiểu dại của chúng. Sự tryền nhiễm TDV chỉ có thể xảy ra trên lý thuyết khi (1) vectơ đưa vào vacxin với độ chuẩn virut đủ để gây nhiễm cho ruột giữa của muỗi, (2) virut này có khả năng sao chép trong biểu mô ruột giữa và có khả năng phát tán tiếp theo ra khỏi ruột giữa, và (3) virut đã được phát tán có thể sao chép trong tuyến nước bọt và lượng virut được nhỏ ra trong nước bọt đủ để truyền nhiễm. Ngưỡng nhiễm virut ở người cần thiết để gây nhiễm cho muỗi đã được thiết lập một cách thích đáng, nhưng nhiễm virut ở người có thể là 10^6-10^8 liều nhiễm ở muỗi $_{50}$ (MID_{50})/ml sau khi gây nhiễm DENV kiểu dại có trong tự nhiên. MID_{50} này dựa trên sự chủng ngừa IT trực tiếp đối với muỗi bằng huyết tương người đã được pha loãng. Phân tích TDV ở động vật linh trưởng không phải người chỉ ra rằng độ chuẩn nhiễm virut sau khi gây miễn dịch TDV là rất thấp (nhỏ hơn 2,4 $\log_{10}$ pfu/ml) và kéo dài trong 2-7 ngày. Với mức độ nhiễm virut thấp và khả năng gây nhiễm, phát tán và truyền nhiễm thấp ở muỗi của TDV, không chắc là các virut vacxin này có thể được truyền nhiễm bởi muỗi trong tự nhiên hoặc gây ra sự nhiễm virut.

Vì vậy, dự định rằng bất kỳ giống cấy chuyển nào của typ huyết thanh (P1-P10) bất kỳ có thể được sử dụng trong chế phẩm để tạo ra vacxin an toàn và hiệu quả chống lại một, hai, ba hoặc cả bốn typ huyết thanh virut sốt xuất huyết.

Tính độc thần kinh ở chuột nhắt đang bú mẹ

Virut vacxin loại nghiên cứu ban đầu được giảm độc lực hết sức đối với tính độc thần kinh ở chuột nhắt ICR mới sinh được nuôi trong nhà ở DVBD/CDC. Toàn bộ chuột nhắt này sống sót sau khi được thử thách ic (trong não - intracerebral) với 10^4 pfu mỗi loại virut dùng làm vacxin. Typ huyết thanh virut sốt xuất huyết -2 16681 kiểu dại, mặt khác, tạo ra tỷ lệ tử vong 62,5% - 100% ở chuột nhắt CDC-ICR này trong các thí nghiệm khác nhau. Trong một số thí nghiệm, chuột nhắt ICR

thương mại do Taconic Labs (Taconic-ICR) cung cấp được sử dụng để nghiên cứu tính độc thần kinh ở chuột nhắt mới sinh. Đã quan sát thấy rằng chuột nhắt Taconic-ICR mới sinh dễ bị nhiễm typ huyết thanhvirut sốt xuất huyết-2 đáng kể hơn so với chuột nhắt CDC-ICR trước đó. Fig. 5A tóm tắt tính độc thần kinh của typ huyết thanh virut sốt xuất huyết-2 16681 kiểu dại ở đàn CDC-ICR và chuột nhắt mới sinh Taconic-ICR được cho thử trong hộp sọ với 10^4 pfu virut. Chuột nhắt Taconic-ICR (tỷ lệ tử vong 100% ở 32 chuột nhắt, thời gian sống sót trung bình 8,3 ± 0,5 ngày) nhạy cảm hơn với việc thử thách typ huyết thanh virut sốt xuất huyết-2 16681 dùng trong não so với chuột nhắt CDC-ICR trước đó (ảnh hưởng nguy hại 91% ở 72 chuột nhắt, thời gian sống sót trung bình là 14,6 ± 2,3 ngày).

Trong các phương pháp khác được nêu làm ví dụ, để đánh giá tính độc thần kinh của TDV MVS, chuột nhắt Taconic-ICR ban đầu được cho thử thách trong não (ic - intracerebrally) với một liều khoảng 10^4 pfu typ huyết thanh virut sốt xuất huyết-2 16681 kiểu dại, virut D2 PDK-53 VV45R, D2/3-V, hoặc TDV 1-4 trong một (n=16) hoặc hai (n=31-32) thí nghiệm (Fig. 5B). Ở liều lượng này, virut loại nghiên cứu D2/3-V, cũng như TDV-1, and TDV-3 MVS thể hiện kiểu hình tính độc thần kinh giảm độc lực hoàn toàn (không ốm hoặc tử vong). Như được mong đợi, typ huyết thanh virut sốt xuất huyết -2 kiểu dại đã được nhận thấy là "gây nguy hại", với thời gian sống sót trung bình của chuột nhắt (average mouse survival time - AST) là 8,3 ± 0,8 ngày. Ở các chuột nhắt Taconic-ICR nhạy cảm với typ huyết thanh virut sốt xuất huyết-2 này, virut loại nghiên cứu D2 PDK-53-VV45R gây ra tỷ lệ tử vong 81,3%. TDV-2 MVS và TDV-4 MVS gây nguy hại không đổi cho chuột nhắt Taconic-ICR, thể hiện giá trị AST là 9,8 ± 1,7, 10,2 ± 1,4, và 11,3 ± 0,4 ngày, tương ứng.

Theo một số phương pháp được nêu làm ví dụ, tính độc thần kinh của virut typ huyết thanh virut sốt xuất huyết-2 16681 kiểu dại được so sánh với tính độc thần kinh của virut D2 PDK-53 VV45R, TDV-2 MVS và TDV-4 MVS, cũng như virut loại nghiên cứu D2/4-V, ở liều lượng thấp hơn 10 lần (10^3 pfu, Fig. 5C). Typ huyết thanh virut sốt xuất huyết-2 kiểu dại vẫn giữ được kiểu hình tính độc thần kinh nguy hại không đổi, với AST là 9,0 ± 1,4 ngày, ở liều lượng thử thách thấp hơn này. 4

virut khác thể hiện kiểu hình tính độc thần kinh trung gian, và mức độ tính độc thần kinh là đặc hiệu typ huyết thanh. Virut D2 PDK-53-VV45R và virut cùng nguồn gốc TDV-2 MVS của nó thể hiện sự giảm độc lực đáng kể (tỷ lệ sống sót 32,3% với AST là 13,1 ± 3,8 ngày và tỷ lệ sống sót 31,2% vớ AST là 10,5 ± 3,4 ngày, tương ứng). Cả TDV-4 MVS và virut D2/4-V loại nghiên cứu đều được giảm độc lực ở mức cao đối với tính độc thần kinh (tỷ lệ sống sót 81,3% với AST bằng 18,8 ± 5,8 ngày và tỷ lệ sống sót là 100%, tương ứng). Kết quả này đã gợi ý rằng MVS của TDV-1 và -3 thể hiện sự giảm độc lực hoàn toàn đối với tính độc thần kinh, trong khi TDV-2 và -4 MVS không giữ được kiểu hình giảm độc lực mà nó gần giống với các ứng viên vacxin virut loại nghiên cứu tương đồng của chúng.

Fig. 5A-5C là đồ thị được nêu làm ví dụ minh họa tính độc thần kinh ở chuột nhắt mới sinh được thử nghiệm với các chế phẩm khác nhau bao gồm typ huyết thanh virut sốt xuất huyết-2 kiểu dại và các virut sốt xuất huyết giảm độc lực khác. Kết quả chúng của nhiều thí nghiệm mô tả vắn tắt tính độc thần kinh của virut typ huyết thanh virut sốt xuất huyết-2 16681 ở chuột nhắt mới sinh CDC-ICR (n=72) và Taconic-ICR (n=32) được cho thử trong hộp sọ với 10^4 pfu virut (A). Tính độc thần kinh của TDV MVS được thử nghiệm ở chuột nhắt Taconic-ICR với liều lượng 10^4 pfu (B) hoặc 10^3 pfu (C). Số động vật được thử nghiệm trên một nhóm ở một thí nghiệm (n=16) hoặc hai thí nghiệm gộp lại (n=31 hoặc 32) được chỉ ra.

Kiểu hình vết tan của WVS và BVS

Các nghiên cứu nhất định được thực hiện để so sánh kiểu hình vết tan của WVS và BVS vớ MVS, các virut sốt xuất huyết kiểu dại và các thể khảm của chúng loại nghiên cứu bắt nguồn từ chủng tương đồng trong phòng thí nghiệm ở các tế bào Vero (Fig. 6). Kích thước vết tan trung bình được tính toán từ 10 vết tan đối với mỗi virut vacxin, nhưng từ số lượng DENV-1, -3, và -4 kiểu dại giảm. Tất cả các virut MVS của TDV-1, -2, và -3 tạo ra vết tan nhỏ hơn đáng kể so với các thể tương đồng kiểu dại của chúng và rất giống (chênh lệch trong khoảng 0,4-mm) với các virut tương đồng loại nghiên cứu của chúng ở các tế bào Vero. TDV-4 MVS cũng nhỏ hơn đáng kể so với DENV-4 kiểu dại, nhưng lớn hơn một chút (0,9 mm) so với thể khảm D2/4-V bắt nguồn từ chủng gốc trong phòng thí nghiệm. Ngoại trừ TDV-2, tất

cả các WVS và BVS của TDV-1, -3, -4 vẫn có kích thước vết tan nhỏ hơn đáng kể so với kích thước vết tan được tạo ra từ các chủng tương đồng kiểu dại của chúng. TDV-2 WVS và BVS tạo ra vết tan tương tự với vết tan của virut DENV-2 kiểu dại ở các tế bào Vero, nhưng khi được thử nghiệm ở các tế bào LLC-MK_2 tất cả các hạt TDV-2 được sản xuất ra tạo ra các vết tan nhỏ hơn một chút so với các vết tan của DENV-2 kiểu dại (1,4 ± 0,4) và tương tự với D2 PDK-53-VV45R được tạo ra trong phòng thí nghiệm (1,0 ± 0,3) (Fig. 6).

Đánh giá các dấu chuẩn kiểu hình giảm độc lực ở virut, bao gồm kiểu hình vết tan nhỏ, độ nhạy với nhiệt độ, sự sao chép giảm ở các tế bào muỗi, sự gây nhiễm/phát tán/truyền nhiễm giảm bởi muỗi và tính độc thần kinh giảm ở chuột nhắt ICR mới sinh, được đánh giá đối với các chế phẩm chứa chủng gốc MVS. Kết quả chỉ ra rằng tất cả các TDV vẫn có kiểu hình giảm độc lực mong đợi giống với các virut vacxin loại nghiên cứu ban đầu. Với những đột biến chịu trách nhiệm cho sự giảm độc lực được bảo toàn ở tất cả MVS, WVS và BV, có thể mong đợi là các kiểu hình giảm độc lực được giữ lại trong nguyên liệu được sản xuất để thử nghiệm lâm sàng trên người.

Fig. 6 là biểu đồ tần số được nêu làm ví dụ minh họa kích thước vết tan của TDV MVS, WVS, và BVS. Đường kính vết tan trung bình ± SD (khoảng sai số) của vết tan virut ở tế bào Vero hoặc LLC-MK_2 trong điều kiện phủ agaroza được đo ở ngày 9 sau gây nhiễm. Các DENV kiểu dại và các virut ứng viên dùng làm vacxin loại nghiên cứu đã được công bố trước đó được dùng làm mẫu đối chứng và so sánh.

Sao chép virut ở các tế bào C6/36 ở muỗi

Các nghiên cứu trước đây đã chứng tỏ rằng virut vacxin khảm trên cơ sở PDK-53 loại nghiên cứu vẫn có kiểu hình giảm độc lực của virut DENV-2 PDK53 bộ khung ở các tế bào C3/36. Trong một số phương pháp được nêu làm ví dụ, TDV MSV, WVS, và BVS được sinh trưởng ở các tế bào C6/36 để xác nhận sự giữ lại của gen đánh dấu giảm độc lực in vitro của chúng sau khi sản xuất ở quy mô lớn. So với virut sốt xuất huyết kiểu dại, ngoại trừ đối với TDV-3, các hạt đã sản xuất thể hiện sự giảm sinh trưởng đáng kể (giảm ít nhất 3 $\log_{10}$ PFU/ml) ở các tế bào C6/36 ở ngày 6 sau gây nhiễm (Fig. 7). Các hạt TDV-3 MSV cũng thể hiện sự sinh trưởng giảm so

với DENV-3 16562 kiểu dại, nhưng sự giảm này là không đáng kể như vậy (giảm 1-2 log_{10} PFU/ml). Tuy nhiên, độ chuẩn của các lô hạt TDV-3 là tương tự (chênh lệch trong khoảng 1 log_{10} PFU/ml) với virut vacxin D2/3-V khảm loại nghiên cứu ban đầu.

Fig. 8 là biểu đồ được nêu làm ví dụ biểu thị sự sinh trưởng hạn chế của TDV MVS, WVS, và BVS ở các tế bào C6/36. Độ chuẩn trung bình ± SD (khoảng sai số) của virut đã được sao chép trong các tế bào C6/36 7 ngày sau gây nhiễm. Virut sốt xuất huyết kiểu hoang và các virut ứng viên dùng làm vacxin loại nghiên cứu đã được công bố trước đó được dùng để so sánh.

Tính độc thần kinh ở chuột nhắt đang bú mẹ

Các thí nghiệm khác được thực hiện để phân tích tính độc thần kinh ở chuột nhắt ICR mới sinh. Liều lượng trong não 10^4 PFU, tỷ lệ sống sót đối với DENV-2 16681 kiểu dại và D2 PDK-53-VV45R là 0% và 188%, tương ứng (Fig. 9A) ở chuột nhắt ICR, nhưng bằng khoảng 20% đối với DENV-2 16681 kiểu dại và 100% đối với D2 PDK-53-VV45R ở chuột nhắt CDC ICR. Trong nghiên cứu này, TDV-1 và TDV-3 MVS được giảm độc lực (tỷ lệ sống sót 100%) đối với chuột nhắt ở liều lượng 10^4 PFU, nhưng MVS của TDV-2 và TDV-4 gây ra tỷ lệ tử vong 100% ở liều lượng trên 10^4 PFU (Fig. 5A). Tuy nhiên, khi được thử nghiệm ở liều lượng 10^3 PFU virut, TDV-2 (tỷ lệ sống sót 31,3%) và TDV-4 (tỷ lệ sống sót 81,3%) thể hiện tính độc thần kinh giảm so với typ huyết thanh virut sốt xuất huyết-2 16681 kiểu dại (tỷ lệ sống sót 0%), và tỷ lệ sống sót của chúng tương tự với tỷ lệ sống sót của ứng viên vacxin loại nghiên cứu D2 PKD-53-VV45R (32,3%) và D2/4-V (100%), tương ứng (Fig. 9B). Mặc dù DENV-1, -3, hoặc -4 kiểu dại không được bao hàm để so sánh trong nghiên cứu này, các nghiên cứu trước đây đã chứng tỏ rằng DENV-1 16007 kiểu dại giảm độc lực ở chuột nhắt CDC-ICR bằng cách dùng trong não, trong khi cả kiểu dại DENV-3 16562 và DENV-4 1036 đều có tính độc virut cao (tỷ lệ sống sót 0%) đối với chuột nhắt CDC-ICR. Dường như 3 DENV kiểu dại này sẽ thể hiện tính độc virut lớn hơn hoặc tương tự với tính độc virut ở chuột nhắt Taconic ICR nhạy cảm hơn. Vì vậy, việc đưa các virut sốt xuất huyết kiểu dại này vào để so sánh với các TDV MVS tương đồng của chúng được coi là không đủ thông tin. Nghiên cứu này đã chỉ ra rằng cả 4 TDV MVS và virut vacxin ứng viên có nguồn gốc từ chủng

gốc trong phòng thí nghiệm đều thể hiện kiểu hình giảm độc lực có thể so sánh được ở chuột nhắt so với DENV-2 16681 kiểu dại.

Fig. 9A-9B là các đồ thị được nêu làm ví dụ về dữ liệu tính độc thần kinh của TDV MVS ở chuột nhắt ICR mới sinh. (A) chủng ngừa trong hộp sọ bằng virut ở liều lượng 10^4 PFU. (B) chủng ngừa trong hộp sọ bằng virut ở liều lượng 10^3 PFU

Tất cả các lô hạt TDV được kiểm tra tính đồng nhất, sự vô trùng, và không chứa các chất không mong muốn. Phân tích trình tự bộ gen đầy đủ bộc lộ rằng một đột biến axit amin khác nữa được tạo ra trong TDV-4 MVS, trong khi 3 TDV MVS khác vẫn có trình tự bộ gen liên ứng của các hạt tiền chủng gốc của chúng. Ở các lô WVS, TDV-3 đạt được đột biến axit amin khác và 3 typ huyết thanh còn lại có 2 sự thế axit amin khác, so với các hạt tiền chủng gốc của chúng. Trình tự bộ gen của cả 4 lô BVS là giống hệt với các lô WVS của chúng. Tất cả các hạt từ hạt P2 đến hạt tiền chủng gốc (P7), chỉ 1 hoặc 2 đột biến không câm xuất hiện trong hạt đã nêu. Giữa các hạt tiền chủng gốc và BCS (P10), chỉ quan sát thấy 1 đến 2 sự thế nucleotit, tất cả trong số chúng đều xuất hiện ở NS2A, 4A, hoặc 4B, ngoại trừ sự thay đổi nucleotit do glyxin bảo tồn và alanin ở gốc E-483. Từ hạt P2 đến hạt BVS (P10), 3 đến 8 phần tử thế nucleotit giống hệt trong hạt TDV bất kỳ đã nêu, và chỉ 2 đến 4 trong số các phần tử thế này tạo ra sự thay đổi axit amin. Không đột biến câm nào trong BVS là thuộc vùng 5'- hoặc 3'-NCR mà nó có thể ảnh hưởng đến sự sao chép virut. Kết quả này gợi ý rằng virut TDV cực kỳ ổn định về mặt di truyền trong quá trình sản xuất. Ba locus giảm độc lực DENV-2 PDK-53 xác định nằm trong 5'NCR, NS1-53, và NS3-250 vẫn không đổi trong trình tự bộ gen liên ứng trong khi cấy chuyển theo chuỗi TDV để tạo ra chủng gốc BVS. TaqMAMA có độ nhạy cao của locus 5'-NCR-57 thể hiện đột biến ngược tối thiểu hoặc không thể phát hiện được ở MVS, WVS, và BVS của TDV. Tỷ lệ đột biến ngược cao nhất là 0,21% được nhận biết ở TDV-2 BVS. Tỷ lệ đột biến ngược của BVS tương đương P10 (<0,07% đến 0,21%) thấp hơn đáng kể so với tỷ lệ đột biến ngược được tạo ra ở các ứng viên vacxin khác sau khi cấy chuyển theo chuỗi ở các tế bào Vero (tỷ lệ đột biến ngược 4-74% ở P10). Điều này gợi ý rằng chiến lược sản xuất các hạt TDV ở quy mô lớn này là thành công, về mặt duy trì độ ổn định di truyền và giữ lại được các các gen đánh

dấu giảm độc lực ở các virut vacxin ứng viên.

Do chủng gốc MVS được bộc lộ ở đây sẽ được dùng để sản xuất các lô WVS và BVS trong tương lai, bảng đánh giá kiểu hình giảm độc lực đầy đủ, bao gồm kiểu hình vết tan nhỏ, độ nhạy với nhiệt độ, sự sao chép giảm ở các tế bào muỗi, sự gây nhiễm/phát tán/truyền nhiễm giảm ở muỗi nguyên vẹn, và tính độc thần kinh giảm ở chuột nhắt ICR mới sinh, được thực hiện đối với tất cả các MVS hoặc các chủng gốc thay thế tương đương của chúng. Đối với các chủng gốc WVS và BVS, kích thước vết tan, khả năng gây nhiễm ở các tế bào muỗi, cũng dược thực hiện để xác nhận sự giảm độc lực của chúng. Kết quả đã chỉ ra rằng tất cả các chủng gốc MVS của 4 typ huyết thanh của TDV vẫn giữ được kiểu hình giảm độc lực mong muốn, chẳng hạn như vết tan nhỏ, khả năng sao chép giảm ở các tế bào C6/36, và tính độc thần kinh giảm đối với chuột nhắt, giống như virut văc xin được tạo ra trong phòng thí nghiệm ban đầu (Fig. 6, 8, và 9). Ngoại trừ TDV-4, tất cả các chủng gốc 3 MVS khác của TDV đều TS ở 39°C như được minh họa trên các Fig 3 và 7.

Đối với các chủng gốc WVS và BVS, hai kiểu hình giảm độc lực, vết tan nhỏ và sự sao chép hạn chế ở các tế bào C6/36, được phân tích và xác nhận. Do có sự thay đổi rất ít về mặt di truyền giữa MVS và BVS, mong đợi rằng chúng sẽ vẫn có kiểu hình giảm độc lực như MVS. Ngoài các thí nghiệm được mô tả trong báo cáo này, độ an toàn và tính sinh miễn dịch của TDV sản xuất được ở chuột nhắt Ag129 và động vật linh trưởng không phải người đã được thử nghiệm.

Các phương pháp được đề xuất ở đây để biểu thị quá trình sản xuất các chủng gốc TDV MVS, WVS, và BVS trong các điều kiện cGMP. Chủng gốc BVS được sử dụng để bào chế TDV tứ giá hiện nay trong các đánh giá thử nghiệm lâm sàng trên người. Chiến lược sản xuất duy nhất để tối ưu hóa độ ổn định di truyền và độ an toàn của MVS được sản xuất được đưa ra trong một số phương pháp được nêu làm ví dụ. Do các locus giảm độc lực chính của TDV đã được mô tả đặc trưng kỹ trước đây và thử nghiệm SNP cực nhạy và có thể định lượng, TaqMAMA được phát triển để hợp nhất trình tự bộ gen và TaqMAMA để nhận biết các hạt tiền chủng gốc tối ưu để tạo ra MVS. Đặc tính di truyền và kiểu hình của MVS được phân tích đầy đủ để xác nhận rằng các virut này vẫn có sự giảm độc lực mong muốn đối với độ an toàn của

vacxin. Đây có thể là vacxin virut sống giảm độc lực duy nhất có thể được phân tích một cách hữu hiệu đối với tất cả các locus di truyền giảm độc lực chính trong quá trình sản xuất từ tiền chủng gốc đến chủng gốc BVS. Kết quả được đưa ra ở đây làm ví dụ cho lợi thế về độ an toàn của vacxin ở các vacxin sống giảm độc lực được thiết kế theo chiến lược này.

Ví dụ 3

Theo một phương pháp khác được nêu làm ví dụ, các cấu trúc có thể bao gồm cải biến virut sốt xuất huyết sống giảm độc lực dựa trên typ huyết thanh virut sốt xuất huyết-2 sống giảm độc lực (TDV-2), ngoài một hoặc nhiều cấu trúc mã hóa miền protein không cấu trúc của typ huyết thanh virut sốt xuất huyết-2 và protein cấu trúc từ các typ huyết thanh bất kỳ trong số các typ huyết thanh virut sốt xuất huyết-1 (TDV-1), virut sốt xuất huyết-3 (TDV-3) hoặc virut sốt xuất huyết-4 (TDV-4). Ngoài ra, các cấu trúc riêng rẽ trong chế phẩm sinh miễn dịch được bộc lộ ở đây có thể được tạo ra với tỷ lệ hoặc nồng độ dựa trên PFU của log PFU so với một cấu trúc khác. Chẳng hạn, chế phẩm sinh miễn dịch có thể được bào chế sao cho TDV-1 có mặt ở nồng độ khoảng $2x10^4$ PFU, TDV-2 có mặt ở nồng độ khoảng $5x10^4$ PFU, TDV-3 có mặt ở nồng độ khoảng $1x10^5$ PFU, và TDV-4 có mặt ở nồng độ khoảng $3x10^5$ PFU, như được minh họa trên Fig. 11B. Nồng độ tương ứng của các cấu trúc typ huyết thanh virut sốt xuất huyết khác nhau trong chế phẩm sinh miễn dịch cũng có thể được biểu thị dưới dạng tỷ lệ tính theo số log của PFU, chẳng hạn như khoảng 4:4:5:5, tương ứng với tỷ lệ của TDV-1:TDV-2:TDV-3:TDV-4, như được minh họa trên Fig. 11B. Theo một số phương án, tỷ lệ có thể là khoảng 5:4:5:5 cũng như các tỷ lệ định trước khác ở đó một hoặc nhiều typ huyết thanh virut sốt xuất huyết sống đã được giảm độc lực cao hơn hoặc thấp hơn so với các typ huyết thanh virut sốt xuất huyết khác. Theo một ví dụ, nồng độ tương đối của TDV-2 có thể giảm so với TDV-1, TDV-3, và TDV-4, và điều này có thể tạo ra chế phẩm sinh miễn dịch có đáp ứng cân bằng tốt hơn đối với cả bốn typ huyết thanh virut sốt xuất huyết khi dùng cho đối tượng. Chẳng hạn, TDV-1, TDV-2, TDV-3, và TDV-4, có thể tạo ra kháng thể trung hòa chống lại cả bốn typ huyết thanh virut sốt xuất huyết (virut sốt xuất huyết sống giảm độc lực-2 và thể khảm của virut sốt xuất huyết -1, -3 và -4), và sử dụng virut

sốt xuất huyết-2 làm bộ khung có thể gây ra đáp ứng tế bào CB8+ T phản ứng chéo và đa chức năng đối với protein không cấu trúc của virut sốt xuất huyết.

Theo một phương pháp khác được nêu làm ví dụ, để kiểm tra hiệu lực và khả năng dung nạp của chế phẩm sinh miễn dịch virut sốt xuất huyết được bộc lộ ở đây ở trẻ em (trước đây rất khó để một quần thể gây miễn dịch một cách hữu hiệu chống lại virut sốt xuất huyết và các vacxin trước đây không có hiệu lực ở quần thể này đặc biệt là trẻ em 9 tuổi hoặc nhỏ hơn), thử nghiệm lâm sàng với đối chứng giả dược, mù đôi, ngẫu nhiên, hai phần được thực hiện. Phần 1 được tiến hành theo cách giảm dần theo tuổi, ở đó đối tượng từ các lứa tuổi 21-45 (n = 38), đối tượng từ các lứa tuổi 12-20 (n = 36), đối tượng từ các lứa tuổi 6-11 (n = 38) và đối tượng từ các lứa tuổi 1,5-5 (n = 36) được cho dùng một liều chế phẩm vacxin virut sốt xuất huyết chứa cả bốn cấu trúc TDV-1, TDV-2, TDV-3 và TDV-4 (Fig. 12A). Phần II được tiến hành ở các đối tượng trẻ em khỏe mạnh ở các lứa tuổi 1,5-11 (n = 211) như là phần mở rộng của phần I (Fig. 12B). Liều vacxin được dùng ở ngày 0, và các mẫu mô (*chẳng hạn*, các mẫu máu) được lấy từ đối tượng ở ngày 0 và ở các thời điểm khác nhau trong toàn bộ thử nghiệm để kiểm tra mức độ nhiễm virut trong máu, tính sinh miễn dịch và độ an toàn tổng thể (các Fig. 12A-12B). Đối với phần I, các liều ngẫu nhiên vacxin và giả dược được dùng ở tỷ lệ giữa vacxin và giả dược là khoảng 2:1. Đối với phần II, các liều ngẫu nhiên vacxin và giả dược được dùng ở tỷ lệ giữa vacxin và giả dược là khoảng 3:1. Trong chế phẩm sinh miễn dịch, TDV-1 có mặt ở nồng độ khoảng $2x10^4$ PFU, TDV-2 có mặt ở nồng độ khoảng $5x10^4$ PFU, TDV-3 có mặt ở nồng độ khoảng $1x10^5$ PFU, và TDV-4 có mặt ở nồng độ khoảng $3x10^5$ PFU.

Theo một phương pháp khác, để nghiên cứu khả năng dung nạp của chế phẩm sinh miễn dịch chống lại virut sốt xuất huyết được bộc lộ ở đây ở trẻ em, tỷ lệ mắc phải các tác dụng bất lợi (adverse effect - AE) được nghiên cứu (dữ liệu không được chỉ ra). Tỷ lệ mắc phải chứng đau, ngứa, và ban đỏ ở vị trí tiêm, cũng như phù, được xác định giữa việc dùng TDV và giả dược (tự báo cáo). Các AE khác nhau được nghiên cứu trong vòng 14 ngày sau khi dùng lần thứ nhất TDV (n = 249) hoặc giả dược (n = 111), và trong vòng 14 ngày sau khi dùng lần thứ hai TDV (n = 249) hoặc giả dược (n = 111). Nhìn chung, không có báo cáo AE nghiêm trọng nào liên quan

đến vacxin trong số các đối tượng tham gia nghiên cứu, và cũng không có sự thay đổi đáng kể nào về mặt hóa học hoặc huyết học của máu, cho thấy chế phẩm vacxin virut sốt xuất huyết theo sáng chế được dung nạp tốt ở trẻ em (dữ liệu không được chỉ ra).

Theo một phương pháp khác được nêu làm ví dụ, để kiểm tra hiệu lực của chế phẩm sinh miễn dịch theo sáng chế ở trẻ em, sự có mặt của kháng thể trung hòa được sản sinh bởi đối tượng sau khi được tiêm chế phẩm vacxin TDV được xác định ở các thời điểm khác nhau lên tới 720 ngày sau tiêm bằng cách sử dụng thử nghiệm vi trung hòa. Như được minh họa trên Fig. 13A-13B, việc sử dụng vacxin TDV gây ra đáp ứng kháng thể trung hòa mà tiếp tục tồn tại cho đến ngày 720, bất kể trạng thái miễn dịch ban đầu của đối tượng (dương tính huyết thanh ở mức đường nền được minh họa trên Fig. 13A; chưa từng tiếp xúc với virut sốt xuất huyết ở mức đường nền được minh họa trên Fig. 13B). Mức kháng thể Denv-1 được biểu thị bởi các đường nét đứt ngắn được đánh dấu bằng các hình kim cương màu đen đặc; mức kháng thể Denv-2 được biểu thị bởi các đường sẫm màu nét liền được đánh dấu bằng các hình vuông màu đen đặc; mức kháng thể Denv-3 được biểu thị bởi các đường nét đứt dài được đánh dấu bằng các hình tam giác màu đen đặc; và mức kháng thể Denv-4 được biểu thị bởi các đường chấm được đánh dấu bằng các chữ X. Các liều được nhận diện bởi mũi tên đặc chỉ lên trên tương ứng với các ngày 0 và 90.

Theo một ví dụ khác, để kiểm tra xem liệu chế phẩm sinh miễn dịch virut sốt xuất huyết có thể gây ra các đáp ứng tứ giá ở trẻ em bất kể trạng thái miễn dịch ban đầu, tính dương tính huyết thanh của đối tượng nghiên cứu được xác định sau khi được tiêm chế phẩm vacxin TDV ở các thời điểm khác nhau lên tới 720 ngày sau tiêm. Như được minh họa trên Fig. 14, lớn hơn khoảng 60% đối tượng nghiên cứu là dương tính huyết thanh ở tất cả các thời điểm sau tiêm được kiểm tra đối với tất cả các typ huyết thanh virut sốt xuất huyết. Lớn hơn khoảng 80% đối tượng là dương tính huyết thanh ở tất cả các thời điểm sau tiêm được thử nghiệm đối với ba hoặc nhiều hơn ba typ huyết thanh virut sốt xuất huyết. Và gần 100% đối tượng dương tính huyết thanh ở tất cả các thời điểm sau tiêm được thử nghiệm đối với hai hoặc nhiều hơn hai typ huyết thanh virut sốt xuất huyết. Dữ liệu đối với các ngày 0, 28, 90

và 120 được thu nhận từ các phần I và II của thử nghiệm lâm sàng (n = 249), trong khi dữ liệu đối với các ngày 180, 360 và 720 được thu nhận từ phần I của thử nghiệm lâm sàng (n = 90). Độ dương tính huyết thanh được xác định là độ chuẩn $MNT_{50} \geq 10$.

Theo một phương pháp khác được nêu làm ví dụ, các thử nghiệm được thực hiện để xác định xem liệu chế phẩm sinh miễn dịch chống lại virut sốt xuất huyết được bộc lộ ở đây có khả năng gây ra đáp ứng miễn dịch đối với cả bốn typ huyết thanh virut sốt xuất huyết ở trẻ em mà âm tính huyết thanh (chưa bao giờ tiếp xúc với virut sốt xuất huyết) ở mức đường nên hay không. Như được minh họa trên Fig. 15, lớn hơn khoảng 40% đối tượng nghiên cứu là dương tính huyết thanh ở tất cả các thời điểm sau tiêm được kiểm tra đối với tất cả các typ huyết thanh virut sốt xuất huyết. Lớn hơn khoảng 60% đối tượng là dương tính huyết thanh ở tất cả các thời điểm sau tiêm được thử nghiệm đối với ba hoặc nhiều hơn ba typ huyết thanh virut sốt xuất huyết. Và gần 100% đối tượng dương tính huyết thanh ở tất cả các thời điểm sau tiêm được thử nghiệm đối với hai hoặc nhiều hơn hai typ huyết thanh virut sốt xuất huyết. Dữ liệu đối với các ngày 0, 28, 90 và 120 được thu nhận từ các phần I và II của thử nghiệm lâm sàng (n = 133), trong khi dữ liệu đối với các ngày 180, 360 và 720 được thu nhận từ phần I của thử nghiệm lâm sàng (n = 40). Độ dương tính huyết thanh được xác định là độ chuẩn $MNT_{50} \geq 10$. Lưu ý rằng chế phẩm tứ giá được bộc lộ ở đây có khả năng gây ra đáp ứng miễn dịch đối với cả bốn typ huyết thanh virut sốt xuất huyết mà có khả năng tiếp tục tồn tại cho đến 720 ngày sau khi dùng.

Ví dụ 4

Theo một phương pháp, các thành viên được ghi danh; 1794 thành viên nhận hai liều TDV được tiêm dưới da ở các tháng 0 và 3 (n=200; Nhóm 1); một liều ở tháng 0 (n=398; Nhóm 2); một liều ở tháng 0 và liều tăng cường ở tháng 12 (n=998; Nhóm 3); hoặc giả dược (n=198; Nhóm 4). TDV tạo ra kháng thể trung hòa chống lại tất cả các DENV, mà nó đạt đỉnh ở tháng 1 và vẫn giữ ở mức cao trên giá trị đường nền ở tháng 6. GMT tháng 6 (95% CI) ở các nhóm 1–4, tương ứng là: 489 (321, 746), 434 (306, 615), 532 (384, 738), 62 (32, 120) đối với DENV-1; 1565 (1145, 2140), 1638 (1286, 2088), 1288 (1031, 1610), 86 (44, 169) đối với DENV-2;

160 (104, 248), 151 (106, 214), 173 (124, 240), 40 (23, 71) đối với DENV-3; và 117 (79, 175), 110 (80, 149), 93 (69, 125), 24 (15, 38) đối với DENV-4. Trong số các thành viên âm tính huyết thanh ban đầu, hai liều TDV gây ra GMT trung bình và tỷ lệ chuyển dạng huyết thanh thành DENV-3 và -4 cao hơn một liều, và tỷ lệ chuyển dạng huyết thanh tứ giá ở tháng 6 là 85,0%, so với 67,6% sau một liều TDV.

Thiết kế nghiên cứu và các thành viên tham gia nghiên cứu

Các nghiên cứu này là nghiên cứu đối chứng giả dược, mù đôi, ngẫu nhiên, đa tâm được thực hiện ở ba bệnh viện/phòng khám ở Panama, Philippin và Cộng hòa Đô-mi-ni-ca. Các thành viên khỏe mạnh độ tuổi 2–17 tuổi được ghi danh vào nghiên cứu và được đánh giá tính thích hợp, sau đó được sắp xếp ngẫu nhiên theo tỷ lệ 1:2:5:1 để nhận TDV theo hai liều ở tháng 0 và tháng 3 (Nhóm 1); một liều ở tháng 0 (Nhóm 2); một liều ở tháng 0 và một liều tăng cường ở tháng 12 (Nhóm 3); hoặc giả dược (Nhóm 4). Tỷ lệ sắp xếp ngẫu nhiên 1:2:5:1 được lựa chọn để tạo ra dữ liệu trên lịch trình một liều, cùng hoặc không cùng liều tăng cường, để sử dụng một cách tiềm năng phác đồ liều lượng này trong quá trình phát triển ở giai đoạn 3. Các nghiên cứu trước đây của TDV đã thiết lập độ an toàn và tính sinh miễn dịch của lịch trình hai liều 2 (tháng 0-tháng 3) và nhóm này (Nhóm 1) được dùng để so sánh với các nhóm một liều lớn hơn nhóm 1 (Nhóm 2 và Nhóm 3).

Quy trình

Thành phần typ huyết thanh của TDV là $2,5 \times 10^4$, $6,3 \times 10^3$, $3,2 \times 10^4$ và $4,0 \times 10^5$ đơn vị hình thành vết tan (PFU) của TDV-1, TDV-2, TDV-3 và TDV-4, tương ứng, trong chế phẩm đông khô. Giả dược là muối đệm phosphat; TDV được hoàn nguyên trong nước để tiêm ở thời điểm sử dụng, và 0,5 mL TDV hoặc giả dược được tiêm dưới da (chẳng hạn vùng cơ delta). Các mẫu được làm đông khô trước đó trong chế phẩm làm ổn định.

Các mẫu máu để xác định kháng thể trung hòa được thu nhận từ các thành viên trong phân nhóm sinh miễn dịch ở các tháng 0 và 3 trước khi nghiên cứu việc sử dụng để điều trị, và ở các tháng 1 và 6, và được phân tích một cách thiết yếu.

Một tiêu chí là tính sinh miễn dịch được minh họa bởi độ chuẩn trung bình

hình học (geometric mean titres - GMT) của các kháng thể trung hòa đối với mỗi trong số bốn typ huyết thanh DENV ở các tháng 1, 3 và 6 bằng cách sử dụng thử nghiệm vi trung hòa [MNT_{50}]. Tiêu chí sinh miễn dịch thứ hai là tỷ lệ dương tính huyết thanh đối với mỗi trong số bốn typ huyết thanh DENV (ở đó dương tính huyết thanh bởi MNT_{50} được định nghĩa là độ chuẩn trung hòa nghịch đảo $\geq$ 10) ở các tháng 1, 3, và 6. Các tiêu chí sinh miễn dịch được tóm tắt theo bộ quy trình (per protocol set - PPS), bao gồm tất cả các thành viên từ tập hợp con sinh miễn dịch mà không vi phạm quy trình chính và các mẫu máu hợp lệ trước và sau khi dùng liều của các thành viên này là sẵn có.

Phân tích thống kê

Thử nghiệm này được thiết kế chủ yếu mang tính mô tả và không dựa trên các giả thuyết vô hiệu chính thức thử nghiệm. Vì vậy, kích cỡ mẫu không được xác định dựa trên tính toán năng lực thống kê chính thức. Cỡ mẫu theo kế hoạch gồm 1800 thành viên (600 trong tập hợp con sinh miễn dịch) được giả thiết để tạo ra số lượng thành viên hợp lý để đánh giá sự tiếp tục tồn tại của đáp ứng miễn dịch sau khi dùng liều đơn và ảnh hưởng của liều tăng cường, và để tạo ra cơ sở dữ liệu độ an toàn thích đáng để giúp cho thử nghiệm hiệu lực giai đoạn 3 ở khoảng 20000 thành viên.

Tỷ lệ dương tính huyết thanh và GMT của kháng thể trung hòa virut sốt xuất huyết được tính toán với khoảng tin cậy (CI) 95% đối với mỗi trong số bốn typ huyết thanh TDV riêng rẽ ở mức đường nền và các tháng 1, 3 và 6. Tỷ lệ phần trăm thành viên dương tính huyết thanh với ít nhất vacxin nhị giá, tam giá, và tứ giá được tóm tắt theo nhóm ở mỗi lần thăm khám trong nghiên cứu. Dữ liệu này cũng được biểu thị bởi trạng thái huyết thanh virut sốt xuất huyết đường nền; tính dương tính huyết thanh được định nghĩa là MNT_{50} nghịch đảo $\geq$ 10 đối với $\geq$1 typ huyết thanh DENV. Dữ liệu độ an toàn được tóm tắt sinh động, với các AE được quan tâm được biểu thị theo tuổi (< 6 và $\geq$ 6 tuổi).

Ở tháng thứ 6, 1743 thành viên được nhận phác đồ theo kế hoạch. Tuổi trung bình của các thành viên nghiên cứu là 7,3 tuổi (dữ liệu không được chỉ ra). Dữ liệu nhân khẩu học đối với PPS (tập hợp con sinh miễn dịch) thường tương ứng với các dữ liệu về tập hợp an toàn, chỉ khác là tỷ lệ từ 2–5 tuổi trong PPS nhỏ hơn trong tập

hợp an toàn. Tỷ lệ thành viên dương tính huyết thanh đối với DENV bất kỳ ở mức đường nền trong PPS là tương tự giữa các nhóm nghiên cứu (54,7% tổng số; nằm trong khoảng từ 51,2–57,4%). TDV gây ra kháng thể trung hòa chống lại cả 4 typ huyết thanh virut sốt xuất huyết ở các thành viên được chủng ngừa. Mức kháng thể trung hòa virut sốt xuất huyết cao nhất được quan sát đối với DENV-2 (Fig. 2), sau đó là DENV-1, và các mức thấp hơn được quan sát đối với DENV-3 và DENV-4. GMT đối với mỗi typ huyết thanh đạt đỉnh ở tháng 1 và vẫn duy trì cao hơn đường nền ở tháng 6 (Ngày 180). GMT tháng 6 (CI 95%) ở các nhóm 1–4, tương ứng, là DENV-1: 489 (321, 746), 434 (306, 615), 532 (384, 738), 62 (32, 120); DENV-2: 1565 (1145, 2140), 1638 (1286, 2088), 1288 (1031, 1610), 86 (44, 169); DENV-3: 160 (104, 248), 151 (106, 214), 173 (124, 240), 40 (23, 71); và DENV-4: 117 (79, 175), 110 (80, 149), 93 (69, 125), 24 (15, 38). Ở các thành viên mà âm tính huyết thanh ở mức đường nền, lịch trình hai liều gây ra GMT DEN-3 và DEN-4 ở tháng 6 (Nhóm 1) cao hơn liều đơn được dùng cho thành viên ở các nhóm 2 và 3, mặc dù CI là chồng chéo (dữ liệu không được chỉ ra). Ở các thành viên mà dương tính huyết thanh ở mức đường nền, đáp ứng GMT là giống nhau ở tất cả các nhóm TDV (dữ liệu không được chỉ ra).

Tỷ lệ dương tính huyết thanh đối với mỗi DENV trong số các thành viên được chủng ngừa TDV tăng đến 87,3–100% ở tất cả các nhóm nghiên cứu ở tháng 1, và vẫn cao đối với mỗi DENV ở tháng 6 (85,0–100%; Fig. 3, bảng bên trái). Ở những thành viên mà âm tính huyết thanh ở đường nền, lịch trình hai liều tạo ra tỷ lệ dương tính huyết thanh đối với DENV-3 (97,5%) và DENV-4 (87,5%) ở tháng 6 cao hơn so với lịch trình một liều (85,7% và 85,3% đối với DENV-3 và 81,4% và 69,3% đối với DENV-4 – dữ liệu không được chỉ ra). Tuy nhiên, các CI 95% là chồng chéo.

Ở đường nền, 44,7% toàn bộ thành viên dương tính với bốn typ huyết thanh với DENV. Trước tháng 1, nhiều hơn 80% thành viên đã được chủng ngừa bằng TDV ở mỗi nhóm nghiên cứu dương tính với bốn typ huyết thanh, và 96% hoặc nhiều hơn tính dương tính với ít nhất ba typ huyết thanh (dữ liệu không được chỉ ra). Tỷ lệ dương tính nhiều typ huyết thanh được duy trì ở tháng 6, với tỷ lệ ở tháng thứ 6 cao hơn một chút sau khi dùng hai liều TDV (dữ liệu không được chỉ ra). Ở các

thành viên âm tính huyết thanh ở đường nền, tỷ lệ dương tính với bốn typ huyết thanh ở tháng thứ 6 ở những người nhận hai liều (85% đối với nhóm 1) cao hơn ở những người nhận một liều (70% và 65,3% đối với các nhóm 2 và 3), mặc dù CI 95% là chồng chéo.

Nhìn chung, 1402 thành viên nhận một liều TDV, và 194 nhận hai liều. Không SAE nào trong số 40 SAE được báo cáo bởi 32 thành viên (nghĩa là 1,8% tập hợp an toàn) có liên quan đến vacxin hoặc quy trình nghiên cứu. Ba SAE dẫn đến rút bỏ nghiên cứu (phản ứng dị ứng với chất màu thực phẩm, ban xuất huyết giảm tiểu cầu miễn dịch, và viêm thận tiểu cầu cấp tính). Một trường hợp tử vong không liên quan đến vacxin hoặc quy trình nghiên cứu ảy ra sau khi kết thúc phân tích 6 tháng (do viêm phổi, lao phổi và sốc nhiễm khuẩn). Hai trường hợp mang thai dẫn đến đình chỉ nghiên cứu, nhưng các thành viên sau đó sinh sản bình thường và các con của họ đều khỏe mạnh.

Không quan sát thấy sự khác biệt về tỷ lệ AE được tự nguyện gửi đi giữa TDV và giả dược, cả sau khi dùng liều thứ nhất so với dùng liều thứ hai, hoặc liên quan đến tính dương tính huyết thanh ở đường nền, hầu hết đều không liên quan đến sự chủng ngừa nghiên cứu (bảng 3). Nhìn chung, 15 trên 562 thành viên (2,7%) trong tập hợp con sinh miễn dịch đã báo cáo các AE được gửi đi một cách tự nguyện mà có liên quan đến vacxin. Các AE được gửi đi một cách tự nguyện là nhẹ ở 161 trong 186 (86,6%) thành viên. Trong số các thành viên dương tính huyết thanh ở đường nền, 5 trong 250 (2,0%) và không thành viên nào trong số 45 (0%) được báo cáo các AE được gửi đi một cách tự nguyện có liên quan đến vacxin sau khi tiêm TDV liều thứ nhất và thứ hai, tương ứng, so với 1 trong số 47 (2,1%) và 2 trong số 252 (0,8%) thành viên nhận giả dược. Trong số các thành viên âm tính huyết thanh ban đầu, các AE được gửi đi một cách tự nguyện có liên quan đến vacxin được báo cáo là 4 trong số 202 (2,0%) và 1 trong số 42 (2,4%) sau khi tiêm TDV liều thứ nhất và thứ hai, tương ứng, so với không thành viên nào trong số 43 thành viên (0%) và 2 trong số 203 thành viên (1,0%) nhận giả dược.

Tính dương tính huyết thanh sau khi chủng ngừa có thể là một tiêu chuẩn quan trọng của hiệu suất vacxin, ở chỗ nó tạo ra bằng chứng về đáp ứng có thể đo

được đối với việc chủng ngừa. Do không có mặt thể tương liên của quá trình bảo vệ, đương nhiên không thể nói cường độ đáp ứng như thế nào là cần thiết để bảo vệ được. Tuy nhiên, một vacxin tạo ra sự miễn dịch thể dịch và miễn dịch tế bào và biểu hiện chuyển dạng huyết thanh có thể xác định được đối với tất cả các typ huyết thanh virut sốt xuất huyết ở đa số cá thể, thậm chí các cá thể mà không tiếp xúc trước đó với virut sốt xuất huyết, gợi ý rằng nó thích hợp để đánh giá trong thử nghiệm hiệu lực vacxin ở quy mô lớn và dường như là một vacxin hữu hiệu chống lại virut sốt xuất huyết ở trẻ em và thanh niên. Phác đồ liều được lựa chọn đối với thử nghiệm này cần tạo ra đáp ứng đa giá ở tỷ lệ lớn nhất các đối tượng âm tính huyết thanh ban đầu.

Trong nghiên cứu ở quy mô lớn được nêu làm ví dụ này, nhóm giai đoạn 2 được lấy ra từ hai vùng địa phương mắc virut sốt xuất huyết (Châu Á và Châu Mỹ La tinh) và tính áng chừng dân số thế giới thực mà sẽ được chủng ngừa bằng TDV. Trong nghiên cứu thứ nhất này để đánh giá một liều so với hai liều TDV được dùng cách nhau 3 tháng, đã chứng tỏ rằng tính sinh miễn dịch thể dịch giảm vẫn mạnh 6 tháng sau liều dùng ban đầu, chỉ giảm nhẹ ở các cá thể được chủng ngừa âm tính huyết thanh ban đầu, nhưng dùng liều thứ hai giúp làm giảm bớt sự giảm này (*chẳng hạn* DENV-4) và làm tăng tỷ lệ của đối tượng đáp ứng miễn dịch với việc chủng ngừa (*chẳng hạn* DENV-3). Không có mối bận tâm nào mới xuất hiện sau 6 tháng trong số 1596 thành viên được chủng ngừa TDV. TDV được khẳng định là an toàn và dung nạp tốt từ lứa tuổi 2 tuổi ở trẻ em và thanh niên, bất kể tình trạng huyết thanh đối với virut sốt xuất huyết vào lúc chủng ngừa.

Mặc dù các vacxin khác chứa thể khảm flavivirut được cấp đăng ký ở một số vùng, các vacxin này không được phê chuẩn đối với trẻ em nhỏ hơn 9 tuổi. Các thể khảm flavivirut này gây ra 35–50% mức bảo vệ đặc hiệu typ huyết thanh đối với DENV-1 và -2, và hiệu lực thấp hơn ở những người nhận mà chưa từng tiếp xúc với virut sốt xuất huyết. Vì vậy, vẫn có nhu cầu đối với vacxin mà an toàn và hiệu quả chống lại tất cả bốn typ huyết thanh DENV ở người nhận thuộc tất cả các lứa tuổi, đặc biệt là người nhận trẻ hơn 9 tuổi, bất kể là đã tiếp xúc với virut sốt xuất huyết và lây nhiễm typ huyết thanh trước đó. Kết quả giai đoạn 2 đầy triển vọng của TDV

giúp khởi đầu đánh giá giai đoạn 3 của phác đồ hai liều trong nghiên cứu được thiết kế để giúp sử dụng TDV trong một khoảng tuổi rộng, bất kể là dương tính huyết thanh virut sốt xuất huyết ở thời điểm chủng ngừa (đặc biệt là ở trẻ em nhỏ tuổi).

Vật liệu và phương pháp

Virut và tế bào

DENV-1 16007, DENV-2 16681, DENV-3 16562, và DENV-4 1034 đóng vai trò làm mẫu đối chứng DENV kiểu dại (wt), và chúng là các virut kiểu gen bố mẹ để dùng cho bốn ứng viên vacxin TDV tái tổ hợp. Các virut loại nghiên cứu nguyên bản TDV, được gọi là D2/1-V, D2 PDK-53-VV45R, D2/3-V, và D2/4-V, được điều chế và mô tả đặc trưng trước đây. Tế bào Vero (thận khỉ xanh châu Phi) được sử dụng để tạo ra ngân hàng tế bào chủng giống gốc và giống sản xuất để sản xuất vacxin thu được từ dòng tế bào CCL81 thuộc Bộ sưu tập chủng giống kiểu Mỹ (American Type Culture Collection (ATCC)) đã được mô tả đặc trưng bởi Tổ chức sức khỏe thế giới (World Health Organization (WHO)) để sản xuất vacxin (WCB-tế bào Vero).

Tạo dẫn xuất của virut TDV tái tổ hợp từ các dòng ADN bổ trợ

Để tạo dẫn xuất lại các virut vacxin ứng viên trong các điều kiện sản xuất cGMP, các dòng ADN bổ trợ lây nhiễm DENV đã được tạo ra trước đó, pD2-PDK-53-VV45R, pD2/1-V, pD2/4-V, và pD2/3-V đã được nối *in vitro* chứa ADN bổ trợ virut với chiều dài bộ gen đầy đủ được sử dụng để tạo ra các sản phẩm phiên mã ARN virut mới bằng cách phiên mã *in vitro* như được mô tả trước đây. Tóm lại, ADN bổ trợ bộ gen DENV đã được duỗi thẳng nhờ XbaI được xử lý với proteinaza K, được chiết bằng phenol/cloroform và được kết tủa trong etanol để loại bỏ protein dư bất kỳ, và sau đó được tạo huyền phù trong dung dịch đệm Tris-EDTA không chứa RNaza trước khi phiên mã. Sự phiên mã *in vitro* được thực hiện bằng cách sử dụng kit phiên mã hiệu suất cao AmpliScribe T7 (Epicentre Technologies) theo quy trình được khuyến cáo bởi nhà sản xuất. Chất tương tự mũ A ARN, m7G(5')ppp(5')A (New England BioLabs), được kết hợp trong suốt 2 giờ phản ứng phiên mã để bổ sung mũ A ở đầu tận cùng 5' vào sản phẩm phiên mã ARN. Tiếp theo các mẫu được xử lý với DNaza I để phân giải ADN bổ trợ khuôn mẫu, sau đó

chiết bằng phenol/cloroform ở độ pH thấp và kết tủa bằng etanol để loại bỏ ADN và protein còn dư. Sản phẩm phiên mã ARN đã tinh chế, được tạo huyền phù trong nước không chứa RNaza, được phân phối trong các phần phân ước 20-µl và được bảo quản ở -80°C cho đến khi sẵn sàng để chuyển nhiễm tế bào. Tính nguyên vẹn và nồng độ của các sản phẩm phiên mã ARN được phân tích bằng cách điện di trên gel agaroza. Mỗi phần phân ước 20-µl được ước tính để chứa ARN virut bộ gen với chiều dài đủ để cho phép chuyển nhiễm 0,4-1 x 10^7 tế bào Vero được chứng nhận dùng trong sản xuất bằng điện chuyển.

Sự chuyển nhiễm của mỗi sản phẩm phiên mã ARN vào tế bào WCB-Vero được thực hiện trong điều kiện cGMP ở Shantha Biotechnics. Các sản phẩm phiên mã ARN của TDV được làm tan giá, được trộn với 400 µl dịch huyền phù tế bào Vero (1 x 10^7 tế bào/ml), và được chuyển đến cuvet dùng trong điện di vô trùng đã được làm lạnh từ trước (khe 4-mm) để điện di bằng hệ thống Gene Pulser Xcell total system (BioRad Laboratories). Mỗi mẫu được tạo xung một lần ở 250V/∞ Ohms/500 µF, được ủ trong 10-15 phút ở nhiệt độ trong phòng, được chuyển đến đĩa 75-cm^2 chứa 30 ml môi trường sinh trưởng tế bào (MEM với 10% FBS), và được ủ ở 36°C ± 1°C, 5% CO_2 trong từ 6 đến 11 ngày. Môi trường nuôi cấy được thu nhận, được làm trong bằng cách ly tâm, được làm ổn định, và bảo quản trong các phần phân ước nhỏ dưới -60°C. Độ chuẩn virut của các dung dịch gốc vacxin ứng viên (được gọi là P1 đối với mức cấy chuyển 1) thu được từ quá trình chuyển nhiễm được xác định bằng thử nghiệm chuẩn độ vết tan ở các tế bào Vero và được sử dụng để nhân giống tiếp các hạt TDV.

Sản xuất các hạt virut TDV

Các hạt virut P1 được sử dụng để nhân giống các lô hạt virut tiền chủng gốc TDV, chủng gốc, chủng sản xuất, và các hạt virut với lượng lớn nhờ một chiến lược được thiết kế để đảm bảo độ ổn định di truyền và độ an toàn tối ưu của các lô sản xuất. Chiến lược này bao gồm ba quá trình tinh chế vết tan theo chuỗi, cũng như phân tích bộ gen của virut ở các mức cấy chuyển khác nhau để chọn lọc quần thể virut tạo dòng tối ưu để tiếp tục sản xuất hạt virut (bảng 1). Tóm lại, các hạt P1 được thu nhận từ tế bào đã chuyển nhiễm được khuếch đại một lần bằng cách gây nhiễm tế

bào Vero ở MOI bằng 0,001 để tạo ra hạt P2. Các phần phân ước của chủng gốc hạt P2 được đánh giá bởi hình thái học vết tan và giải trình tự toàn bộ bộ gen virut. Chủng gốc P2 đã được xác nhận về mặt di truyền được đặt trên các lớp đơn tế bào Vero với môi trường phủ lên trên như được mô tả trong phần chuẩn độ vết tan dưới đây để tạo ra vết tan được phân lập tốt. Sau khi làm hiện màu bằng đỏ trung tính, sáu vết tan riêng rẽ từ mỗi trong số 4 typ huyết thanh của các virut vacxin được phân lập (các dòng vết tan từ A đến F) và được trộn trong 0,5 ml môi trường nuôi cấy (cấy chuyển P3). Mỗi trong số sáu dịch huyền phù chứa vết tan được tham gia hai vòng tinh chế vết tan nữa, tạo ra các hạt virut được tinh chế hai lần và ba lần từ vết tan ở các giống cấy chuyển P4 và P5, tương ứng. Virut P5 được khuếch đại qua hai lần cấy chuyển Vero liên tiếp để tạo ra chủng gốc hạt P7.

Phân tích di truyền ba locus giảm độc lực TDV chính bằng cách sử dụng phương pháp đọc trình tự điểm và/hoặc thử nghiệm đột biến khuếch đại không ghép đôi trên cơ sở Taqman (Taqman-based mismatched mutation assay - TaqMAMA) như đã bộc lộ trước đây, và phân tích kiểu hình vết tan được thực hiện để sàng lọc toàn bộ 24 hạt P7. Các hạt có đặc tính ban đầu thích hợp sau đó được mô tả đặc trưng bằng cách giải trình tự bộ gen đầy đủ. Kết quả của những phân tích này, một trong số 6 hạt P7 (dòng A-F) của mỗi typ huyết thanh TDV được lựa chọn làm hạt tiền chủng gốc, dựa trên sự có mặt của DENV-2 PDK-53 đột biến giảm độc lực, biến đổi trình tự bộ gen tối thiểu, và kiểu hình vết tan mong muốn. Mỗi hạt tiền chủng gốc được lựa chọn được phát triển thành hạt virut chủng gốc (MVS hoặc P8) bằng quy trình cấy chuyển một lần virut ở MOI bằng 0,001 trong nhiều bình 175 cm^2 chứa tế bào Vero. Ngoại trừ TDV-4 MVS, các hạt virut chủng gốc được thu nhận ở 8-10 ngày sau gây nhiễm (post infection - pi). Giống gốc MVS được thu nhận ở 6-10 ngày sau gây nhiễm (pi), được làm trong bằng cách ly tâm, được làm ổn định bằng cách bổ sung dung dịch sucroza/phosphat/glutamat (nồng độ cuối 7,5 % sucroza, 3,4 mM kali dihydro phosphat, 7,2 mM dikali hydro phosphat, 5,4 mM mononatri glutamat, tương ứng) và 0,95 đến 1,90% FBS (nồng độ cuối). TDV-4 MVS được bào chế theo cách khác để tối ưu hóa hiệu suất của nó. Một cách ngắn gọn, nhiều bình chứa tế bào đã được gây nhiễm với hạt tiền chủng gốc TDV-4 ở MOI bằng 0,001 với sự có mặt

của 0,1% F-127™, poloxamer 407, (các copolyme khối EO-PO khác đã được đánh giá và có thể thay thế ở đây, xem patent đã được phát hành) mà đã được chứng minh là tăng cường độ ổn định nhiệt của virut DENV. Môi trường gây nhiễm được thu nhận ở các ngày 6-10 sau khi gây nhiễm, và được làm ổn định với 17% FBS (nồng độ cuối), được gom lại, và làm lạnh đông. Cả bốn dung dịch gốc TDV MVS được bảo quản dưới dạng các phần phân ước 1-ml ở dưới -60°C.

Các hạt virut sản xuất TDV (working virut seeds - WVS) được bào chế bằng cách cấy chuyển một lần MVS ở MOI bằng 0,001 trong dịch nuôi cấy tế bào Vero. Các quy trình tương tự với quy trình sản xuất MVS, chỉ khác là chúng được nuôi cấy trong các nhà máy tế bào nhiều lớp (6360 cm^2). Các dung dịch gốc WVS được lọc qua màng lọc 10 µM và 0,45µM, được làm ổn định với cùng chất làm ổn định được dùng cho MVS, được phân thành các phần phân ước vào các chai PETG 30ml hoặc các lọ lạnh 2,0 ml, và được bảo quản ở dưới -60°C.

Theo các phương pháp nhất định, các hạt virut với lượng lớn (BVS) được sản xuất bằng cách gây nhiễm nhiều nhà máy tế bào (6360 cm^2 mỗi nhà máy) chứa tế bào Vero nhập dòng bằng 90 mL WVS đã pha loãng để đạt được MOI bằng 0,001. Môi trường được sử dụng để pha loãng dịch chủng ngừa WVS chứa 0,1% F-127™ không chứa huyết thanh. Sau khi hấp phụ trong 1,5 giờ, các tế bào được rửa 3 lần bằng PBS, và 800 ml môi trường DMEM không chứa huyết thanh được bổ sung vào mỗi nhà máy tế bào, và các nhà máy được ủ ở 36(±1)°C ở 5(±0,5)% CO_2. Sau khi ủ trong bốn ngày, các phần phân ướ nhỏ của môi trường được gom để kiểm tra độ vô trùng. Virut được thu nhận giữa ngày 5 và ngày 10 sau gây nhiễm, và được làm trong ngay lập tức bằng cách lọc qua màng lọc cỡ lỗ 0,45 um, và mỗi phần gom 1L virut đã lọc trong được làm ổn định bằng cách bổ sung 500 ml dung dịch đệm 3x FTA (nồng độ cuối là 15% trehaloza, 1,0% Pluronic® F-127™ poloxamer 407, 0,1% albumin người USP trong PBS, pH 7,4). Virut đã được làm ổn định được phân phối vào các chai PETG 1-L và được bảo quản đông lạnh dưới -60°C để kết hợp và kiểm tra giám sát chất lượng tiếp theo. Tất cả các dịch thu nhận virut đã được làm ổn định với độ chuẩn virut trên 10^5 PFU/ml và mức ADN còn dư chấp nhận được được làm tan giá nhanh trong chậu nước ở 32°C, sau đó được gom và kết hợp một cách vô

trùng. Mỗi BVS đơn giá được gom lại được phân phối vào các đồ chứa PETG đã được dán nhãn và được bảo quản ở dưới -60°C cho đến khi sử dụng tiếp theo.

Kiểm tra chất lượng sản phẩm trong quá trình sản xuất

Các hạt MVS, WVS, và BVS được kiểm tra về tính đồng nhất, tính vô trùng, và các chất ngẫu nhiên có thể phát hiện được. Tính đồng nhất của dung dịch vacxin gốc được xác nhận bằng RT-PCR với các đoạn mồi đặc hiệu typ huyết thanh TDV. Các mảnh ADN bổ trợ được khuếch đại chứa vị trí nối thể khảm E/NS1 để cho phép nhận biết mỗi trong số bốn typ huyết thanh TDV. Mỗi hạt được thử nghiệm trong tất cả 4 phản ứng RT-PCR đặc hiệu typ huyết thanh để xác nhận tính đồng nhất của virut và không bị nhiễm tạp chéo các typ huyết thanh TDV khác loại. Thử nghiệm tính vô trùng được thực hiện theo USP 71 (United States Pharmacopeia, section 71). Thử nghiệm mycoplasma được thực hiện.

Các thử nghiệm *in vitro* và *in vivo* sau đây về sự nhiễm tạp virut đều được thực hiện bằng cách sử dụng các dịch thu nhận TDV chưa được làm trong, chưa được làm ổn định được thu hồi trong quá trình sản xuất hạt virut. Môi trường gây nhiễm đã thu nhận trước tiên được trung hòa với kháng huyết thanh đa dòng ở thỏ DENV (Inviragen) ở 36 ± 1°C trong 1 giờ để làm bất hoạt DENV. Đối với thử nghiệm *in vitro*, các hạt virut đã trung hòa được chủng vào ba dòng tế bào chỉ thị, MRC5, VERO và MA104, trong các bình 25 cm^2. Virut Echo (đối chứng CPE) hoặc virut quai bị (đối chứng hấp phụ hem) được sử dụng làm mẫu đối chứng CPE hoặc hấp phụ hem dương tính, tương ứng. Tất cả các tế bào được giám sát CPE hằng ngày trong tổng số 14 ngày. Ở cuối ngày thứ 14, dịch nổi nuôi cấy được loại bỏ và được thay bằng 10 mL dung dịch tế bào hồng cầu (red blood cell - RBC) chuột lang (3 mL RBC chuột lang 0,5% trong dung dịch muối đệm phosphat, được pha thành 10 mL với môi trường sinh trưởng tế bào). Sau đó, các bình được ủ ở 5 ± 3°C trong 30 phút được theo sau bởi bước ủ ở nhiệt độ trong phòng trong 30 phút. Các lớp đơn được rửa bằng PBS và được quan sát với độ phóng đại 10 X về sự có mặt của các cụm RBC hình sao bất kỳ đối với sự hấp phụ hem.

Thử nghiệm in vivo đối với các chất ngẫu nhiên được thực hiện ở chuột nhắt đang bú mẹ, chuột nhắt sau cai sữa và chuột lang. Chuột nhắt đang bú mẹ được

chủng bằng 0,1ml hoặc 0,01 ml (10 con chuột nhắt trong mỗi nhóm liều lượng) mẫu hạt đã trung hòa bằng kháng huyết thanh DENV bằng cách tiêm trong màng bụng (intraperitoneal - ip). Tương tự, 10 chuột nhắt sau cai sữa, mỗi con được chủng ip bằng 0,5 ml hoặc 0,03 ml mẫu. Mỗi con chuột lang (5 con/nhóm) được chủng ngừa ip bằng 5,0 mL. Chuột nhắt đang bú mẹ được quan sát hằng ngày về tình trạng bệnh tật và tỷ lệ tử vong trong tổng số 14 ngày sau chủng ngừa. Chuột nhắt sau cai sữa được quan sát trong tổng số 28 ngày, và chuột lang được quan sát trong tổng số 42 ngày sau chủng ngừa. Các mẫu thí nghiệm đạt tiêu chí chấp nhận được nếu ≥80% động vật đã chủng ngừa vẫn khỏe mạnh trong suốt giai đoạn quan sát.

Thử nghiệm *in vivo* đối với các chất nhiễm tạp cũng được thực hiện trong trứng gà có phôi và được kiểm soát. Đối với mỗi mẫu, 10 trứng gà mái có phôi (9 ngày tuổi) mỗi trứng được chủng ngừa với 0,5 mL mẫu đã trung hòa bằng kháng huyết thanh DENV vào dịch niệu nang và được ủ ở 35°C trong 3 ngày. Dịch niệu nang từ 10 trứng này được thu hoạch, gom lại và cấy chuyển vào dịch niệu nang của 10 trứng có phôi mới (10-11 ngày tuổi; 0,5mL/trứng) và được ủ ở 35°C trong 3 ngày nữa. Tương tự, đối với mỗi mẫu, 10 trứng có phôi (6-7 ngày tuổi) mỗi trứng được chủng ngừa với 0,5 mL trên một trứng (BVS đơn giá TDV-2) hoặc 0,25 mL trên một trứng (BVS TDV-1, TDV-3 và TDV-4) bằng cách tiêm vào màng phôi và được ủ ở 35°C trong 9 ngày. Các màng phôi từ 10 trứng này được thu hoạch, gom lại và dịch huyền phù 10% được cấy chuyển vào màng phôi của 10 trứng có phôi mới (6-7 ngày tuổi; 0,5 mL/trứng) và được ủ ở 35°C trong 9 ngày nữa. Trứng được chủng ngừa vào dịch niệu nang (cả chủng ngừa ban đầu và chủng ngừa cấy chuyển) được quan sát khả năng sống sót sau 3 ngày ủ. Tất cả các phần gom của dịch niệu nang được kiểm tra hoạt tính ngưng kết hồng cầu bằng cách sử dụng hồng cầu nhóm O của gà con, chuột lang và người ở 4°C và 25°C. Trứng được chủng ngừa vào màng phôi (cả chủng ngừa ban đầu và chủng ngừa cấy chuyển) được quan sát khả năng sống sót sau 9 ngày ủ.

Thử nghiệm vết tan virut và thử nghiệm miễn dịch tập trung

Độ chuẩn virut được xác định bằng thử nghiệm vết tan hoặc thử nghiệm miễn dịch tập trung bằng cách sử dụng tế bào Vero. Thử nghiệm vết tan được thực hiện

trong các lớp phủ kép agaroza bên trên ở các đĩa 6 lỗ chứa tế bào Vero đã nhập dòng như đã mô tả trước đây, và chúng cũng được sử dụng để đánh giá kiểu hình vết tan của hạt TDV. Để so sánh chính xác, kích thước vết tan của tất cả các virut được đo và được so trong cùng một thí nghiệm. Sau khi hiện màu bằng đỏ trung tính ở ngày thứ 9 sau chủng ngừa, lên tới 10 vết tan đã được phân lập tốt đối với mỗi virut được đo để tính kích thước vết tan trung bình. Số vết tan được đo đối với DENV-1, -3, và -4 kiểu dại ít hơn, kích thước vết tan lớn hơn của chúng thường không cho phép đo 10 vết tan được tách riêng hoàn toàn.

Do TDV tứ gái chứa tất cả bốn typ huyết thanh DENV, thử nghiệm miễn dịch tập trung đặc hiệu typ huyết thanh DENV đã được phát triển để định lượng mỗi thành phần TDV trong chế phẩm tứ giá. Thử nghiệm miễn dịch tập trung của mỗi MVS TDV riêng rẽ được so sánh với thử nghiệm vết tan để đảm bảo kết quả chuẩn độ virut là có thể so sánh được giữa hai thử nghiệm. Thử nghiệm miễn dịch tập trung được thực hiện trong các đĩa 6 lỗ chứa tế bào Vero nhập dòng được gây nhiễm các virut đã được pha loãng theo dãy nồng độ. Các tế bào được phủ bên trên bằng một lớp môi trường muối cân bằng (môi trường BSS/YE-LAH) chứa 0,7% carboxymetyl xenluloza độ nhớt cao (Sigma) và được ủ trong 7 ngày ở 37°C với 5% CO_2. Sau khi loại bỏ lớp phủ, các tấm chứa tế bào được rửa 3 lần bằng PBS, được cố định bằng axeton lạnh 80% trong 30 phút ở -20°C, được rửa một lần bằng PBS, và được phong bế bằng dung dịch đệm phong bế chứa 2,5% (w/v) sữa bột không béo, 0,5% Triton X-100, 0,05% Tween-20 trong PBS ở 37°C trong 30 phút. Các tế bào đã phong bế được ủ với MAb đặc hiệu typ huyết thanh DENV đã được pha loãng, 1F1 (DENV-1), 3H5 (DENV-2), 8A-1 (DENV-3), hoặc 1H10 (DENV-4) trong dung dịch đệm phong bế qua đêm ở 37°C trong 1 giờ hoặc 4°C, được rửa 3 lần bằng dung dịch đệm rửa (0,05% Tween-20 trong PBS), và được ủ với IgG dê kháng chuột tinh chế ái lực được tiếp hợp với phosphataza kiềm hoặc peroxidaza từ cây cải ngựa (HRP) (Jackson Immuno Research Laboratories) ở 37°C trong 45-60 phút. Các đĩa được rửa 3 lần trước khi có chất thích hợp, 1-Step NBT/BCIP cùng với chất ức chế (Pierce) đối với phosphataza kiềm hoặc kit Vector-VIP (Vector Labs) đối với HRP, được bổ sung vào để phát triển màu. Việc phát triển màu được dừng bằng cách rửa với nước

khi các tâm đã được phát triển màu hoàn toàn. Các tâm miễn dịch đã nhuộm được hiện màu trực tiếp và được đếm trên hộp sáng.

Trình tự bộ gen

Bộ gen với chiều dài đầy đủ của MVS và WVS được đọc trình tự (xem dưới đây). Tóm lại, ARN virut được chiết từ các hạt TDV bằng cách sử dụng kit ARN virut QIAamp (Qiagen), và các mảnh ADN bổ trợ chồng lặp bao phủ toàn bộ bộ gen được khuếch đại bằng cách sử dụng kit Titan One Tube RT-PCR (Roche Applied Science, Inc.). Các mảnh ADN bổ trợ đã khuếch đại được tinh chế trên gel trước khi phân tích trình tự bằng cả đoạn mồi xuôi và ngược bằng cách sử dụng kit phân tích trình tự theo chu trình BigDye Terminator v3.1 (Applied Biosystems). Sản phẩm phản ứng chuỗi được làm sạch bằng cách sử dụng kit tinh chế BigDye Xterminator (Applied Biosystems), và chạy trên thiết bị phân tích gen 3130xl (Applied Biosystems) ở DVBD/CDC. Phần mềm Lasergene SeqMan (DNAStar, Inc) được sử dụng để phân tích và so sánh bộ gen.

Thử nghiệm đột biến khuếch đại không ghép đôi trên cơ sở Taqman (TaqMAMA)

TaqMAMA là thử nghiệm đa hình nhạy dùng để định lượng một nucleotit đơn lẻ được phát triển để cho phép đánh giá chính xác mức đột biến ngược ở locus giảm độc lực 5'NC-57, và được tối ưu tiếp để dùng cho nghiên cứu này. ARN virut đã được chiết từ MVS và WVS được phân tích bằng TaqMAMA với cả hai bộ đoạn mồi/đầu dò Taqman đặc hiệu đối với kiểu dại hoặc vùng 5'NC-57 trên vacxin. Các đoạn mồi xuôi được sử dụng để phát hiện DENV-2 kiểu dại và các trình tự vacxin là D2-41-GC và D2-40-TT, tương ứng. nucleotit đầu tận cùng 3' của mỗi đoạn mồi xuôi bắt cặp với nucleotit 5'NCR-57 đặc hiệu đối với mỗi virut, trong khi nucleotit liền kề với nucleotit đầu tận cùng 3' trên mỗi đoạn mồi khác với trình tự bộ gen virut DENV-2 để gia tăng hiệu quả bắt cặp. Đoạn mồi ngược, CD-207, và đầu dò Taqman, CD-169F, đối với cả kiểu dại và bộ vacxin đều giống hệt. Các trình tự của các đoạn mồi và đầu dò cũng như các điều kiện chu trình được mô tả trước đây. RT-PCR thời gian thực được thực hiện với hệ thống iQ5 hoặc CFX-95 (BioRad), bằng cách sử dụng kit BioRad iScript RT-PCR (đối với đầu dò), trong phản ứng 25-µl chứa 5 µl khuôn ARN virut, 0,4 uM mỗi đoạn mồi, và 0,2 uM đầu dò. Các phản ứng lặp ba đối

với mỗi thử nghiệm đặc hiệu kiểu dại và vacxin được thực hiện đối với mỗi mẫu. Số lượng bản sao bộ gen được xác định theo đường cong tiêu chuẩn được tạo ra đối với mỗi kiểu gen virut, trong đó các tiêu chuẩn ARN là các sản phẩm phiên mã bắt nguồn từ plasmit chứa nt 1-2670 của mỗi ADN bổ trợ đặc hiệu bộ gen. Ngoài ra, tính đặc hiệu của thử nghiệm được xác nhận bằng cách kiểm tra mỗi tiêu chuẩn ARN với bộ đoạn mồi/đoạn dò kiểu gen tương đồng để đảm bảo tối thiểu hóa khả năng phản ứng chéo trong mỗi thí nghiệm. Các kết quả được báo cáo dưới dạng tỷ lệ phần trăm của bộ gen virut thể hiện đột biến ngược. Trước đây, do khả năng phản ứng chéo lớn hơn mà hạn chế mức ARN đầu vào đối với thử nghiệm này, độ nhạy phát hiện đột biến ngược ban đầu là khoảng 0,1% (khả năng khu biệt). Do đó, thử nghiệm đã được tối ưu hóa hơn bằng cách sử dụng thiết bị PCR thời gian thực cải tiến và các kit phản ứng, và khả năng phản ứng chéo là giảm đáng kể ở mức đầu vào khuôn ARN cao hơn nhiều (7-8 $\log_{10}$ bản sao). Sự tối ưu hóa này dẫn đến sự cải thiện đáng kể độ nhạy phát hiện đột biến ngược, giảm xuống đến 0,01-0,07%.

Sao chép virut ở các tế bào C6/36 ở muỗi và độ nhạy với nhiệt độ ở các tế bào Vero thuộc động vật có vú

Kiểu hình sao chép của bốn dung dịch gốc MVS TDV và các virut DENV-1, -2, -3, và -4 kiểu dại được đánh giá ở các tế bào C6/36 ở muỗi (*Aedes albopictus*). Các tế bào C6/36 được nuôi trên đĩa 6 lỗ được lây nhiễm trong hai mẫu lặp mới mỗi virut ở MOI bằng 0,001 và được ủ với 4 ml/lỗ môi trường DMEM chứa 2% FBS trong thiết bị ủ chứa 5% CO_2 ở 28°C. Các phần phân ước nhỏ của dịch nổi nuôi cấy được thu gom đối với mỗi virut vào ngày thứ sáu sau cấy nhiễm, được trộn với một thể tích giống hệt môi trường chứa 40% FBS, và được bảo quản ở -80°C cho đến khi được đưa đi chuẩn độ vết tan virut.

Việc xác định độ nhạy với nhiệt độ được thực hiện bằng cách so sánh khả năng phát triển của virut ở 39°C so với khả năng phát triển ở 37°C ở 5 ngày sau khi cấy nhiễm tế bào Vero trên đĩa 6 lỗ. Tế bào được gây nhiễm trong mẫu lặp bốn với mỗi virut ở MOI bằng 0,001 ở 37°C. Sau khi hấp phụ virut, dịch nuôi cấy đã cấy nhiễm được ủ với 4 ml/lỗ môi trường DMEM chứa 2% FBS trong 2 tủ ủ chứa 5% CO_2 riêng biệt, một tủ đặt (hai đĩa lặp) ở 37°C và tủ còn lại đặt ở 39°C. Các phần

phân ước (50-µl) dịch nổi nuôi cấy được thu gom vào ngày thứ 5 sau cấy nhiễm, được trộn với một thể tích giống hệt DMEM chứa 40% FBS, và được bảo quản ở -80°C cho đến khi được đưa đi chuẩn độ vết tan virut. Nhiệt độ ủ được hiệu chỉnh với nhiệt kế đã được hiệu chỉnh tại nhà máy có thể truy nguyên NIST (-1 đến 51 °C; ERTCO).

Gây nhiễm, phát tán, và truyền nhiễm ở muỗi

Muỗi *Aedes aegypti* được sử dụng cho nghiên cứu này là từ bầy được xác định năm 2002 từ một ngôi làng gần Mae Sot (16' N, 33' E), Thái lan. Sau khi nở ra từ ấu trùng, muỗi trưởng thành được giữ ở 28°C ở chu kỳ quang học 16:8 (sáng:tối) và cho ăn một cách tùy ý dung dịch 10% sucroza. Muỗi cái năm đến sáu ngày tuổi được sử dụng để cho ăn bột máu đã gây nhiễm hoặc được chủng ngừa trong ngực intrathoracic (intrathoracic - IT). Các phần phân ước của TDV mới nuôi cấy và DENV kiểu dại được sử dụng ngay sau khi thu nhận (không cần bất kỳ chu trình lạnh đông-tan giá nào) để tạo ra bột máu chứa virut như được chỉ ra dưới đây để lây nhiễm qua đường miệng. Dịch nổi virut còn lại được tạo huyền phù với FBS đến nồng độ cuối bằng 20%, và các phần phân ước được bảo quản ở -80°C để chuẩn độ vết tan virut tiếp và dùng cho các thí nghiệm chủng ngừa IT. Các hạt TDV mới được bào chế cho các thí nghiệm này được khuếch đại từ các hạt tiền chủng gốc ở các tế bào Vero, và được xem là các dạng tương đương MVS TDV.

Bột máu đã nhiễm được điều chế bằng cách trộn virut mới tạo ra ở tỷ lệ 1:1 với máu gà con đã được loại fibrin (Colorado Serum Company) vào ngày gây nhiễm qua đường miệng. Muỗi được ngừng cho ăn đường qua đêm và sau đó tạo ra hỗn hợp virut:máu trong 1 giờ bằng cách sử dụng hệ cho ăn qua màng Hemotek (Discovery Workshops). Phần phân ước 50-µl bột máu được giữ ở -80°C để chuẩn độ ngược liều lượng virut. Muỗi cái đã được cho ăn no được phân loại theo cỡ trong điều kiện gây mê lạnh và được đặt trong các hộp bìa cứng để cho ăn tùy thích dung dịch sucroza 10%. Các hộp bìa cứng được đặt ở 28°C với chu kỳ quang học là 16:8 h (sáng:tối). Sau 14 ngày, 25-30 con muỗi từ mỗi nhóm virut được gây mê bằng cách cho tiếp xúc với trietylamin (Flynap ®, Carolina Biological Supply Company) và một chân sau được lấy khỏi và đặt trong 0,5 ml DMEM với 10% FBS và 5%

penicillin/streptomycin (tương ứng 100U/ml và 100µg/ml). Nước bọt được thu gom bằng cách cài một vòi hút của muỗi đã gây mê vào một mao quản chứa dung dịch 2,5% FBS và 25% sucroza. Muỗi được để tiết nước bọt trong ít nhất 15 phút và sau đó các ống mao quản và toàn thân được đặt trong các ống riêng rẽ chứa DMEM. Thân muỗi, chân muỗi và nước bọt được bảo quản ở -80°C cho đến khi chúng được chuẩn độ và đánh giá sự nhiễm virut. Để chủng ngừa IT, muỗi được gây mê lạnh và được chủng bằng khoảng 50 pfu virut trong 0,34 µl chất chủng. Muỗi đã chủng được giữ trong 7 ngày trong cùng điều kiện như đã mô tả trên đây. Sau đó, muỗi được gây mê, và nước bọt và cơ thể của chúng được thu gom như đã mô tả trên đây. Các mẫu được bảo quản ở -80°C cho đến khi xử lý tiếp theo.

Để xử lý các mẫu nhằm chuẩn độ virut, các mẫu cơ thể và chân được đồng hóa với BB được phủ đồng (Crossman Corporation, NY) ở 24 chu trình/giây trong 4 phút bằng cách sử dụng máy nghiền trộn, và sau đó được làm trong bằng cách ly tâm ở 3.000 x g trong 3 phút. Các mẫu nước bọt được ly tâm ở 3.000 x g trong 3 phút để đuổi dịch ra khỏi mao quản. Các dịch pha loãng mười lần của dịch đồng hóa cơ thể và chân và các mẫu nước bọt được kiểm tra sự có mặt của virut lây nhiễm bằng thử nghiệm vết tan. Kết quả từ cơ thể, chân, và nước bọt được sử dụng để xác định tỷ lệ gây nhiễm, phát tán, và truyền nhiễm, tương ứng.

Tính độc thần kinh ở chuột

Chuột nhắt ICR cái đang trong thời kỳ mang thai thu được từ Taconic Labs, và được giám sát vài lần mỗi ngày để xác định thời gian sinh gần đúng của chuột con. Trong một thí nghiệm được đưa ra, khoảng 12-24 giờ sau sinh, hai con chuột con trên tổng số tám con đối với mỗi virut (n=16), được thử thách với 10^3 đến 10^4 pfu virut trong 20 µl dịch pha loãng bằng cách chủng ngừa trong sọ (intracranial - ic) có sử dụng kim tiêm cỡ 30. Các động vật được giám sát ít nhất 3 lần mỗi ngày trong ít nhất 32 ngày sau thử thách. Khi xuất hiện dấu hiệu ốm đầu tiên (lông xù, lưng cong, sút cân, cử động bất thường, liệt, hoặc ngủ lịm) các động vật được gây mê bằng cách gây mê đến chết bằng khí isofluran, sau đó bằng cách làm trật khớp cổ. Ngày gây chết không đau đớn sau gây nhiễm là "thời gian đến khi ốm/tử vong" hoặc "thời gian sống" đối với động vật. Các thí nghiệm trên động vật được thực hiện sau quy trình

động vật được phê chuẩn bởi DVBD/CDC IACUC.

Tạo dẫn xuất của virut hạt chủng gốc

Hạt virut chủng gốc (Master Virut Seed - MVS) TDV-1

Trình tự nucleotit của bộ gen virut khảm và trình tự axit amin đã được suy ra của các protein dịch mã được đưa ra ở đây. Chẳng hạn, trình tự polynucleotit TDV-1 bao gồm các trình tự được biểu thị bởi SEQ ID NO: 1, 3, 5 và 7, và trình tự polypeptit TDV-1 bao gồm các trình tự được biểu thị bởi SEQ ID NO: 2, 4, 6 và 8. Hầu hết gen prM-E (nt 457 đến -2379, gạch chân) là đặc hiệu virut DEN-1 16007 kiểu dại (wt); bộ gen còn lại là đặc hiệu virut DEN-2 PDK-53. Tất cả các dạng thay thế được tạo ra khác với virut kiểu dại (D1 16007 hoặc D2 16681), cũng như các thể siêu đột biến (thay đổi so với dòng ADN bổ trợ được thiết kế) được phát hiện trong MVS được đánh dấu.

Các dạng thay thế được bao hàm trong bộ gen và protein:

Các vị trí nối giữa bộ khung D1 (prM-E) và D2:

a. MluI (nt 451-456): đột biến câm được thiết kế, nt-453 A-to-G

b. NgoMIV (nt 2380-2385): đột biến được thiết kế, nt-2381/2382 TG-to-CC (tạo ra sự thay đổ Val thành Ala ở E-482)

Bộ khung virut D2 PDK-53 (thay đổi so với D2 16681 kiểu dại): tất cả đều in đậm

a. vùng không mã hóa 5'(NCR)-57 (nt-57 C thànhT): locus giảm độc lực chính (màu đỏ)

b. NS1-53 Gly thành Asp (nt-2579 G-to-A): locus giảm độc lực chính (màu đỏ)

c. NS2A-181 Leu thành Phe (nt-4018 C thành T)

d. NS3-250 Glu thành Val (nt-5270 A thành T): locus giảm độc lực chính (màu đỏ)

e. nt-5547 (gen NS3) đột biến câm T thành C

f. NS4A-75 Gly thành Ala (nt-6599 G thành C)

* nt-8571 đột biến câm C thành T của PDK-53 không được thiết kế trong virut vacxin

DEN-1 prM-E (thay đổi từ D1 16007 kiểu dại)

a. đột biến câm nt-1575 T thành C được tạo ra để loại bỏ vị trí XbaI bẩm sinh

Các thay thế khác được tìm thấy trong hạt vacxin (0,03% nt khác với dòng gốc ban đầu)

Hạt virut chủng gốc (Master Virut Seed - MVS) TDV-2

a. NS2A-116 Ile thành Leu (nt-3823 A thành C, in đậm)

b. NS2B-92 Glu-to-Asp (nt-4407 A thành T, in đậm)

c. nt-7311 A thành G đột biến câm (in đậm)

Trình tự nucleotit của bộ gen virut khảm và trình tự axit amin đã được suy ra của các protein dịch mã được đưa ra ở đây. Chẳng hạn, trình tự polynucleotit TDV-2 bao gồm các trình tự được biểu thị bởi SEQ ID NO: 9, 11, 13 và 15, và trình tự polypeptit TDV-2 bao gồm các trình tự được biểu thị bởi SEQ ID NO: 10, 12, 14 và 16. Virut được thiết kế dựa trên virut D2 PDK-53. Tất cả các dạng thay thế được thiết kế khác với virut DEN-2 16681 kiểu dại (cũng như virut bố mẹ đối với PDK-53), cũng như các dạng đột biến thêm (thay đổi từ dòng ADN bổ trợ được thiết kế) được phát hiện trong MVS được đánh dấu.

Các dạng thay thế được bao hàm trong bộ gen và protein:

Bộ khung virut D2 PDK-53 (thay đổi so với D2 16681 kiểu dại): tất cả đều in đậm

a. vùng không mã hóa 5' (NCR)-57 (nt-57 C thành T): locus giảm độc lực chính (màu đỏ)

b. prM-29 Asp thành Val (nt-524 A thành T)

c. nt-2055 C thành T (gen E) đột biến câm

d. NS1-53 Gly thành Asp (nt-2579 G thành A): locus giảm độc lực chính

(màu đỏ)

e. NS2A-181 Leu thành Phe (nt-4018 C thành T)

f. NS3-250 Glu thành Val (nt-5270 A thành T): locus giảm độc lực chính (màu đỏ)

g. nt-5547 (gen NS3) đột biến câm T thành C

h. NS4A-75 Gly thành Ala (nt-6599 G thành C)

* nt-8571 đột biến câm C thành T của PDK-53 không được thiết kế trong virut vacxin

Các gen đánh dấu dòng được thiết kế (đột biến câm):

a. nt-900 T thành C đột biến câm: gen đánh dấu dòng lây nhiễm

Các thay thế khác được tìm thấy trong hạt vacxin (0,02% nt khác với dòng gốc ban đầu

a. prM-52 Lys thành Glu (nt-592 A thành G), in đậm

b. NS5-412 Ile thành Val (nt-8803 A thành G), in đậm

Hạt virut chủng gốc (Master Virut Seed - MVS) TDV-3

Trình tự nucleotit của bộ gen virut khảm và trình tự axit amin đã được suy ra của các protein dịch mã được đưa ra ở đây. Chẳng hạn, trình tự polynucleotit TDV-3 bao gồm các trình tự được biểu thị bởi SEQ ID NO: 17, 19, 21 và 23, và trình tự polypeptit TDV-3 bao gồm các trình tự được biểu thị bởi SEQ ID NO: 18, 20, 22 và 24. Hầu hết gen prM-E (nt -457 đến -2373, gạch chân) là đặc hiệu virut DEN-3 16562 kiểu dại (wt); trình tự nucleotit còn lại là đặc hiệu virut DEN-2 PDK-53. Protein E của virut DEN-3 có ít hơn hai axit amin so với protein E của DEN-2. Vì vậy, vị trí nt bắt đầu từ NgoMIV ít hơn 6 nt so với vị trí nt PDK-53 ở DEN-2 ban đầu. Tất cả các dạng thay thế được tạo ra khác với virut kiểu dại (DEN-3 16562 hoặc DEN-2 16681), cũng như các thể siêu đột biến (thay đổi so với dòng ADN bổ trợ được thiết kế) được đánh dấu.

Các dạng thay thế được bao hàm trong bộ gen và protein:

Vị trí nối:

a. MluI (nt 451-456): đột biến câm được thiết kế, nt-453 A-to-G

b. NgoMIV (nt 2374-2379): đột biến được thiết kế, nt-2375/2376 TG-to-CC (tạo ra sự thay đổi Val thành Ala ở E-480)

Bộ khung virut D2 PDK-53 (thay đổi từ D2 16681 kiểu hoang): in đậm

a. vùng không mã hóa 5'(NCR)-57 (nt-57 C thànhT): locus giảm độc lực chính (màu đỏ)

b. NS1-53 Gly thành Asp (nt-2573 G thành A): locus giảm độc lực chính (màu đỏ)

c. NS2A-181 Leu thành Phe (nt-4012 C thành T)

d. NS3-250 Glu thành Val (nt-5264 A thành T): locus giảm độc lực chính (màu đỏ)

e. Đột biến câm nt-5541 (gen NS3) T thành C

f. NS4A-75 Gly thành Ala (nt-6593 G thành C)

* Đột biến câm nt-8565 C thành T của PDK-53 không được thiết kế trong virut vacxin

Đột biến được thiết kế trong DEN-3 prM-E (thay đổi từ D3 16562 kiểu dại)

a. Đột biến câm nt-552 C thành T được thiết kế: gen đánh dấu dòng

b. E-345 His thành Leu (nt-1970 A thành T) được thiết kế để sao chép một cách hiệu quả trong dịch nuôi cấy

Các thay thế khác được tìm thấy trong hạt vacxin (0,02% nt khác với dòng gốc ban đầu)

a. Đột biến E-223 Thr thành Ser (nt-1603 A thành T, in đậm)

b. Đột biến câm nt-7620 A thành G (in đậm)

Hạt virut chủng gốc (Master Virut Seed - MVS) TDV-4

Trình tự nucleotit của bộ gen virut khảm và trình tự axit amin đã được suy ra của các protein dịch mã được đưa ra ở đây. Chẳng hạn, trình tự polynucleotit TDV-4 bao gồm các trình tự được biểu thị bởi SEQ ID NO: 25, 27, 29 và 31, và trình tự polypeptit TDV-4 bao gồm các trình tự được biểu thị bởi SEQ ID NO: 26, 28, 30 và 32. Hầu hết gen prM-E (nt -457 đến -2379, gạch chân) là đặc hiệu virut DEN-4 1036 kiểu dại (wt); trình tự nucleotit còn lại là đặc hiệu virut DEN-2 PDK-53. Tất cả các dạng thay thế được tạo ra khác với virut kiểu dại (DEN-3 16562 hoặc DEN-2 16681), cũng như các thể siêu đột biến (thay đổi so với dòng ADN bổ trợ được thiết kế) được đánh dấu.

Các dạng thay thế được bao hàm trong bộ gen và protein:

Vị trí nối:

a. MluI (nt 451-456): đột biến câm được thiết kế, nt-453 A-to-G

b. NgoMIV (nt 2380-2385): đột biến được thiết kế, nt-2381/2382 TG-to-CC (tạo ra sự thay đổi Val thành Ala ở E-482)

Bộ khung virut D2 PDK-53 (thay đổi từ D2 16681 kiểu dại)

a. Vùng không mã hóa 5'(NCR)-57 (nt-57 C thànhT): locus giảm độc lực chính (màu đỏ)

b. NS1-53 Gly thành Asp (nt-2579 G thành A): locus giảm độc lực chính (màu đỏ)

c. NS2A-181 Leu thành Phe (nt-4018 C thành T, in đậm)

d. NS3-250 Glu thành Val (nt-5270 A thành T): locus giảm độc lực chính (màu đỏ)

e. Đột biến câm nt-5547 (gen NS3) T thành C (in đậm)

f. NS4A-75 Gly thành Ala (nt-6599 G thành C, in đậm)

* đột biến câm nt-8571 C thành T của PDK-53 không được thiết kế trong virut vacxin

Các dạng thay thế được thiết kế trong dòng ADN bổ trợ

a. C-100 Arg thành Ser (nt-396 A thành C) được thiết kế: có thể cải thiện sự sao chép virut trong dịch nuôi cấy

b. Đột biến câm nt-1401 A thành G được thiết kế

c. E-364 Ala thành Val (nt-2027 C thành T) được thiết kế: có thể cải thiện sự sao chép virut trong dịch nuôi cấy

d. E-447 Met thành Leu (nt-2275 A thành C) được thiết kế: có thể cải thiện sự sao chép virut trong dịch nuôi cấy

Các thay thế khác được tìm thấy trong hạt vacxin (0,06% nt khác với dòng gốc ban đầu)

a. Đột biến câm nt-225 (gen C) A thành T (in đậm)

b. Đột biến NS2A-66 Asp thành Gly (nt-3674 A thành G) (in đậm)

c. NS2A-99 Lys thành hỗn hợp Lys/Arg (nt-3773 A thành hỗn hợp A/G, in đậm)

d. Đột biến câm nt-5391 C thành T (gen NS3) (in đậm)

e. NS4A-21 Ala thành Val (nt-6437 C thành T, in đậm)

f. đột biến câm nt-7026 T thành hỗn hợp C/T (in đậm)

g. đột biến câm nt-9750 A thành C (in đậm)

Tất cả các chế phẩm và phương pháp được bộc lộ và yêu cầu bảo hộ ở đây có thể được tiến hành và thực hiện không cần thực nghiệm quá mức khi xem xét bộc lộ này. Trong khi chế phẩm và phương pháp theo sáng chế đã được mô tả liên quan đến các phương án ưu tiên, người có hiểu biết trung bình trong lĩnh vực sẽ hiểu rõ rằng, một số thay đổi có thể được áp dụng đối với chế phẩm và phương pháp và về các bước hoặc trình tự các bước của phương pháp được mô tả ở đây mà không đi trệch khỏi quan niệm, tinh thần và phạm vi bảo hộ của sáng chế. Cụ thể hơn nữa, rõ ràng là, một số tác nhân mà chúng có quan hệ cả về mặt hóa học và về mặt sinh lý có thể

được thay thế cho các tác nhân được mô tả ở đây trong khi vẫn đạt được các kết quả tương tự hoặc giống hệt. Tất cả các thay thế và cải biến như vậy là hiển nhiên đối với người có hiểu biết trung bình trong lĩnh vực và được cho là thuộc nội dung, tinh thần và phạm vi của sáng chế như được xác định trong phần yêu cầu bảo hộ kèm theo.

YÊU CẦU BẢO HỘ

1. Dược phẩm chứa chế phẩm vacxin tứ giá bao gồm cả bốn typ huyết thanh virut sốt xuất huyết bao gồm virut sốt xuất huyết-2 sống giảm độc lực và thể khảm virut sốt xuất huyết-2/1, virut sốt xuất huyết-2/3, và virut sốt xuất huyết-2/4,

trong đó virut sốt xuất huyết-2 là ở dạng DENV-2 16681 có nguồn gốc từ biến thể DEN-2 PDK-53 với một đột biến ba ở NS1-53, ở 5'NC-57 và ở NS3-250 sao cho vị trí axit amin 250 của protein NS3 chứa một gốc valin, và trong đó thể khảm này có bộ gen DEN-2 PDK-53 đã nêu là bộ khung virut và một hoặc nhiều gen protein cấu trúc mã hóa capsid, tiền màng/màng hoặc lớp vỏ của bộ gen DEN-2 PDK53 đã nêu hoặc tổ hợp của chúng được thay bằng một hoặc nhiều gen protein cấu trúc tương ứng từ DEN-1, DEN-3 hoặc DEN-4, và

trong đó

nồng độ của typ huyết thanh virut sốt xuất huyết-2 sống giảm độc lực trong dược phẩm dưới dạng log đơn vị hình thành vết tan (PFU) thấp hơn ít nhất một nửa so với log PFU của một hoặc nhiều typ huyết thanh virut sốt xuất huyết khác hoặc

nồng độ của thể khảm virut sốt xuất huyết-2/3 dưới dạng PFU lớn hơn ít nhất một nửa so với typ huyết thanh virut sốt xuất huyết-2 trong chế phẩm vacxin tứ giá.

2. Dược phẩm theo điểm 1,

trong đó thể khảm virut sốt xuất huyết-2/1 có hai đột biến nữa sao cho vị trí axit amin 116 của protein NS2A chứa một gốc leuxin và vị trí axit amin 92 của protein NS2B chứa một gốc axit aspartic,

trong đó virut sốt xuất huyết-2 có hai đột biến nữa sao cho vị trí axit amin 52 của protein prM chứa một gốc axit glutamic và vị trí axit amin 412 của protein NS5 chứa một gốc valin,

trong đó thể khảm virut sốt xuất huyết-2/3 có bộ gen DEN-2 PDK-53 đã nêu là bộ khung virut và có gen prM-E (nt-457 đến -2373) của bộ gen DEN-2 PDK-53 đã nêu được thay bằng gen prM-E tương ứng từ DEN-3 16562 kiểu dại và trong đó thể

khảm virut sốt xuất huyết-2/3 này có một đột biến nữa ở gen prM-E tương ứng đã nêu từ DEN-3 16562 kiểu dại sao cho vị trí axit amin 223 của protein E chứa một gốc serin, và

trong đó thể khảm virut sốt xuất huyết-2/4 có hai đột biến nữa sao cho vị trí axit amin 66 của protein NS2A chứa một gốc glyxin và vị trí axit amin 21 của protein NS4A chứa một gốc valin và trong đó thể khảm virut sốt xuất huyết-2/4 là kiểu gen hỗn hợp đối với vị trí axit amin 99 của protein NS2A chứa một gốc arginin hoặc một gốc lysin.

3. Dược phẩm theo điểm 1 hoặc 2,

trong đó thể khảm virut sốt xuất huyết-2/1 có bộ gen DEN-2 PDK-53 đã nêu là bộ khung virut và có gen prM-E (nt-457 đến -2379) của bộ gen DEN-2 PDK-53 đã nêu được thay thế bằng gen prM-E tương ứng từ DEN-1 16007 kiểu dại,

trong đó thể khảm virut sốt xuất huyết-2/3 có bộ gen DEN-2 PDK-53 đã nêu là bộ khung virut và có gen prM-E (nt-457 đến -2373) của bộ gen DEN-2 PDK-53 đã nêu được thay thế bằng gen prM-E tương ứng từ DEN-3 16562 kiểu dại, và

trong đó thể khảm virut sốt xuất huyết-2/4 có bộ gen DEN-2 PDK-53 đã nêu là bộ khung virut và có gen prM-E (nt-457 đến -2379) của DEN-2 PDK-53 đã nêu được thay thế bằng gen prM-E tương ứng từ DEN-4 1036 kiểu dại.

4. Dược phẩm theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó tỷ lệ giữa thể khảm virut sốt xuất huyết-2/1 với typ huyết thanh virut sốt xuất huyết-2 sống giảm độc lực, với thể khảm virut sốt xuất huyết-2/3, với thể khảm virut sốt xuất huyết-2/4 là khoảng 4:4:5:5 log đơn vị hình thành vết tan (plaque forming unit - PFU).

5. Dược phẩm theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó tỷ lệ giữa thể khảm virut sốt xuất huyết-2/1, với typ huyết thanh virut sốt xuất huyết-2 sống giảm độc lực, với thể khảm virut sốt xuất huyết-2/3, với thể khảm virut sốt xuất huyết-2/4 là khoảng 5:4:5:5 log PFU.

6. Dược phẩm theo điểm bất kỳ trong số các điểm từ 1 đến 3 chứa ít nhất một trong số:

thể khảm virut sốt xuất huyết-2/1 có nồng độ từ 1,0 x 10^3 đến 5 x 10^5 PFU; typ huyết thanh virut sốt xuất huyết-2 sống giảm độc lực có nồng độ từ 1,0 x 10^3 đến 5 x 10^5 PFU;

thể khảm virut sốt xuất huyết-2/3 có nồng độ từ 5,0 x 10^3 đến 5 x 10^5 PFU; và

thể khảm virut sốt xuất huyết-2/4 có nồng độ từ 1,0 x 10^4 đến 5 x 10^6 PFU.

7. Dược phẩm theo điểm bất kỳ trong số các điểm từ 1 đến 3 chứa 2,5 x 10^4 PFU virut sốt xuất huyết-2/1, 6,3 x 10^3 PFU virut sốt xuất huyết-2 sống giảm độc lực, 3,2 x 10^4 PFU thể khảm virut sốt xuất huyết-2/3 và 4,0 x 10^5 PFU thể khảm virut sốt xuất huyết-2/4.

8. Dược phẩm theo điểm bất kỳ trong số các điểm nêu trên, trong đó dược phẩm này còn chứa chất đệm làm ổn định để giảm sự thoái hóa của các virut sốt xuất huyết.

9. Dược phẩm theo điểm 8, trong đó chất đệm làm ổn định bao gồm trehaloza và albumin và tùy ý, poloxamer 407.

Fig. 1

Biến thể di truyền trong số các thể khảm D2/4 (so với D2 16681 kiểu dại và D4-1036)

Bộ gen	D2			mối nối	D4				mối nối	D2											
Gen	NCR	C		prM	E					NS1	NS2A			NS3			NS4A		NS4B	NS5	
Kiểu đột biến*	PDK-53	hạt	Eng	MluI	hạt	gen đánh dấu	Eng	Eng	NgoMIV	PDK-53	hạt	hạt	PDK-53	PDK-53	hạt	PDK-53	hạt	PDK-53	hạt	PDK53	hạt
Vị trí NT bộ gen	C57T	A225T	A396C	A453G	C647G/C	A1401G	C2027T	A2275	TG2380/1CC	G2579A	A3674G	A3773A/G	C4018T	A5270T	C5391T	T5547C	C6437T	G6599C	T7026C/T	C8571T**	A9750C
Vị trí AA-protein	NCR	C- cấm	C-R100S	prM-cấm	prM-T79R/T	E- cấm	E-A364V	E-M447L	E-V482A	NS1-G53D	NS2A-D66G	NS2A-K99K/R	NS2A-L181F	NS3-E250V	NS3-cấm	NS3-cấm	NS4A-A21V	NS4A-G75A	NS4B-cấm	NS5- cấm	NS5-cấm
D2-16681	C	A	A (R)	A					TG (V)	G (G0	A (D)	A (K)	C (L)	A (E)	C	T	C (A)	G (G)	T	C	A
D2-PDK-53	T	-	-	|						A (D)	-	-	T (F)	T (V)	-	c	-	C (A)	-	t	-
D4-1036					G (T)	A	C (A)	A (M)													
D2/4V1 được tạo dòng (pD2/4-VP1i	T	-	C (S)	g	-	g	T (V)	C (L)	CC (A)	A (D)	-	-	T (F)	T (V)	-	c	-	C (A)	-	-	-
DENVax-4 (MVS)	T	t	C (S)	g	-	g	T (V)	C (L)	CC (A)	A (D)	G (G)	A/G (K/R)	T (F)	T (V)	t	c	T (V)	C (A)	c/t	-	c

"-" Giống như D2 16681 hoặc D4 1036; chữ cái nt nhỏ: đột biến câm ở vùng đọc mở

*: PDK-53: kiểu gen đặc hiệu D2-PDK-53 (so với 16681), Đỏ: locus PDK-53 giảm độc lực chính; Hạt: Thể đột biến chỉ được phát hiện ở hạt virut đặc hiệu và không được phát hiện ở dòng ban đầu; Eng: đột biến được tạo ra đối với các dòng D2/4; MluI và NgoMV: vị trí RE được tạo ra ở mối nối D2/4

**: C8571T (đột biến câm PDK-53) được bao hàm trong hầu hết các dòng khảm D2/4

Fig. 2

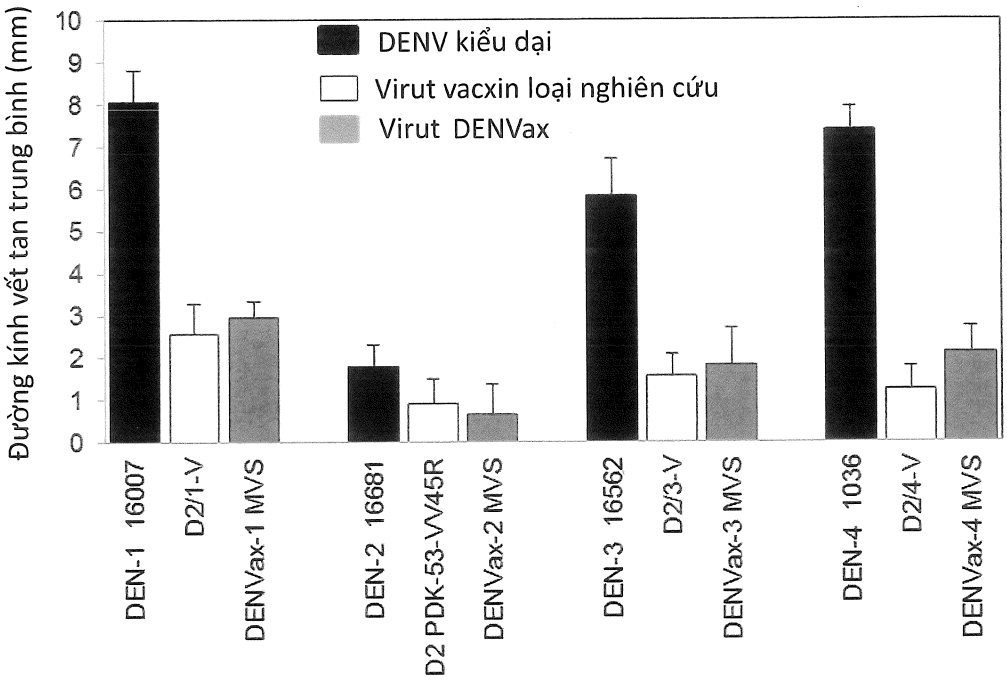

Đường kính vết tan trung bình (mm)
10
9
8
7
6
5
4
3
2
1
0
DENV kiểu dại
Virut vacxin loại nghiên cứu
Virut DENVax
DEN-1 16007
D2/1-V
DENVax-1 MVS
DEN-2 16681
D2 PDK-53-VV45R
DENVax-2 MVS
DEN-3 16562
D2/3-V
DENVax-3 MVS
DEN-4 1036
D2/4-V
DENVax-4 MVS

Fig. 3

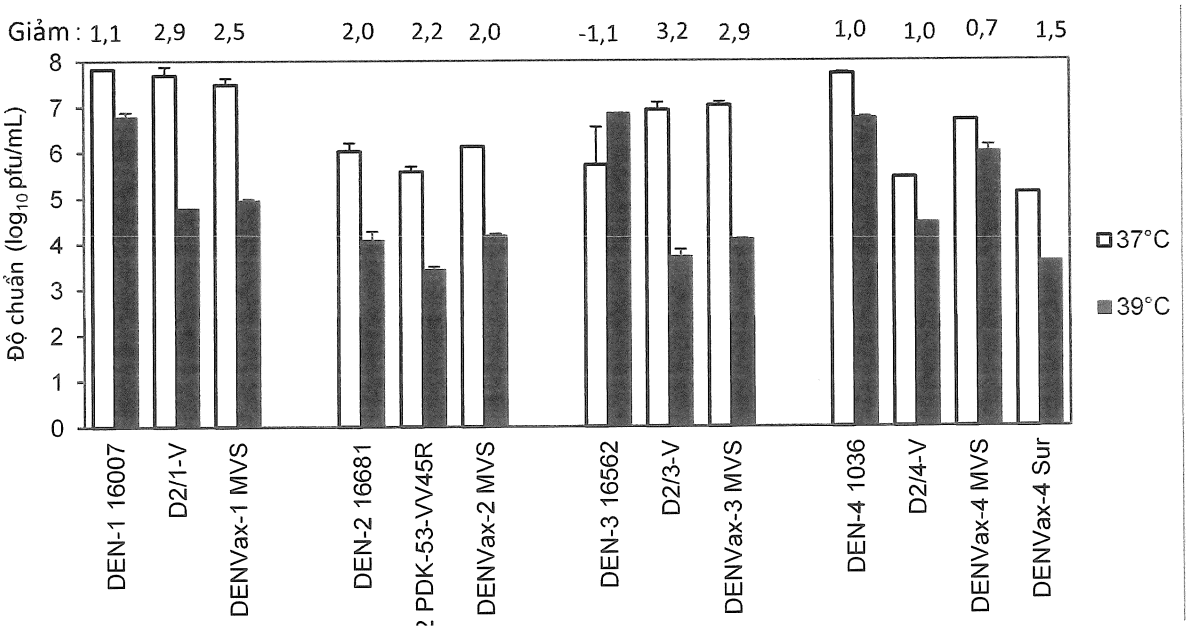

Giảm : 1,1 2,9 2,5 2,0 2,2 2,0 -1,1 3,2 2,9 1,0 1,0 0,7 1,5
Độ chuẩn (log10 pfu/mL)
0
1
2
3
4
5
6
7
8
DEN-1 16007
D2/1-V
DENVax-1 MVS
DEN-2 16681
PDK-53-VV45R
DENVax-2 MVS
DEN-3 16562
D2/3-V
DENVax-3 MVS
DEN-4 1036
D2/4-V
DENVax-4 MVS
DENVax-4 Sur
37°C
39°C

Fig. 4

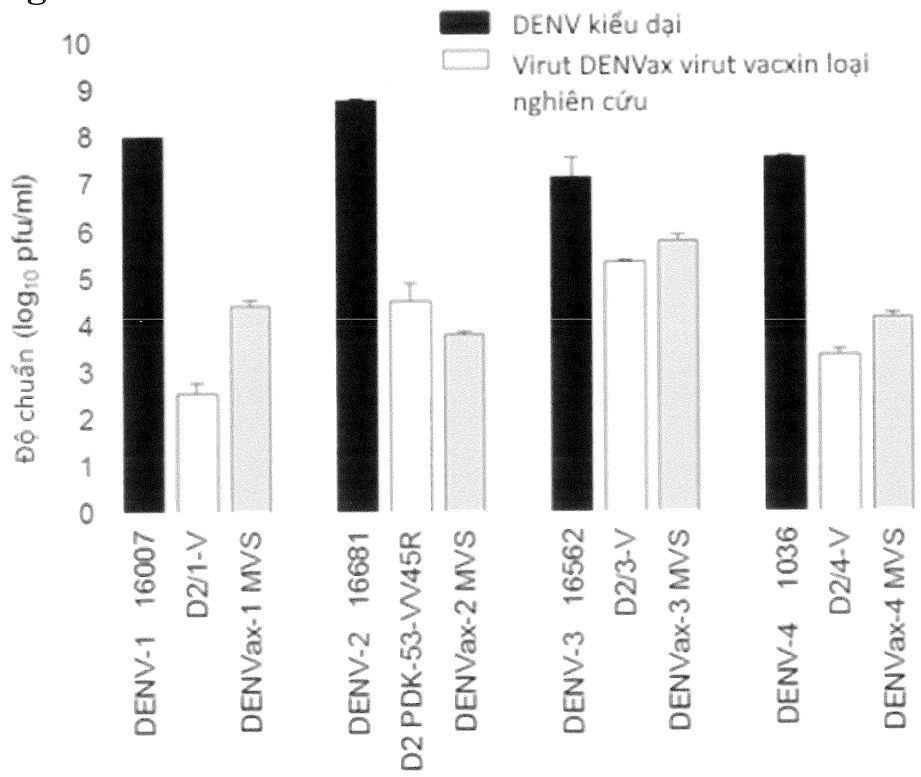

DENV kiểu dại
Virut DENVax virut vacxin loại nghiên cứu
Độ chuẩn (log10 pfu/ml)
10
9
8
7
6
5
4
3
2
1
0
DENV-1 16007
D2/1-V
DENVax-1 MVS
DENV-2 16681
D2 PDK-53-VV45R
DENVax-2 MVS
DENV-3 16562
D2/3-V
DENVax-3 MVS
DENV-4 1036
D2/4-V
DENVax-4 MVS

Fig. 5A-5C

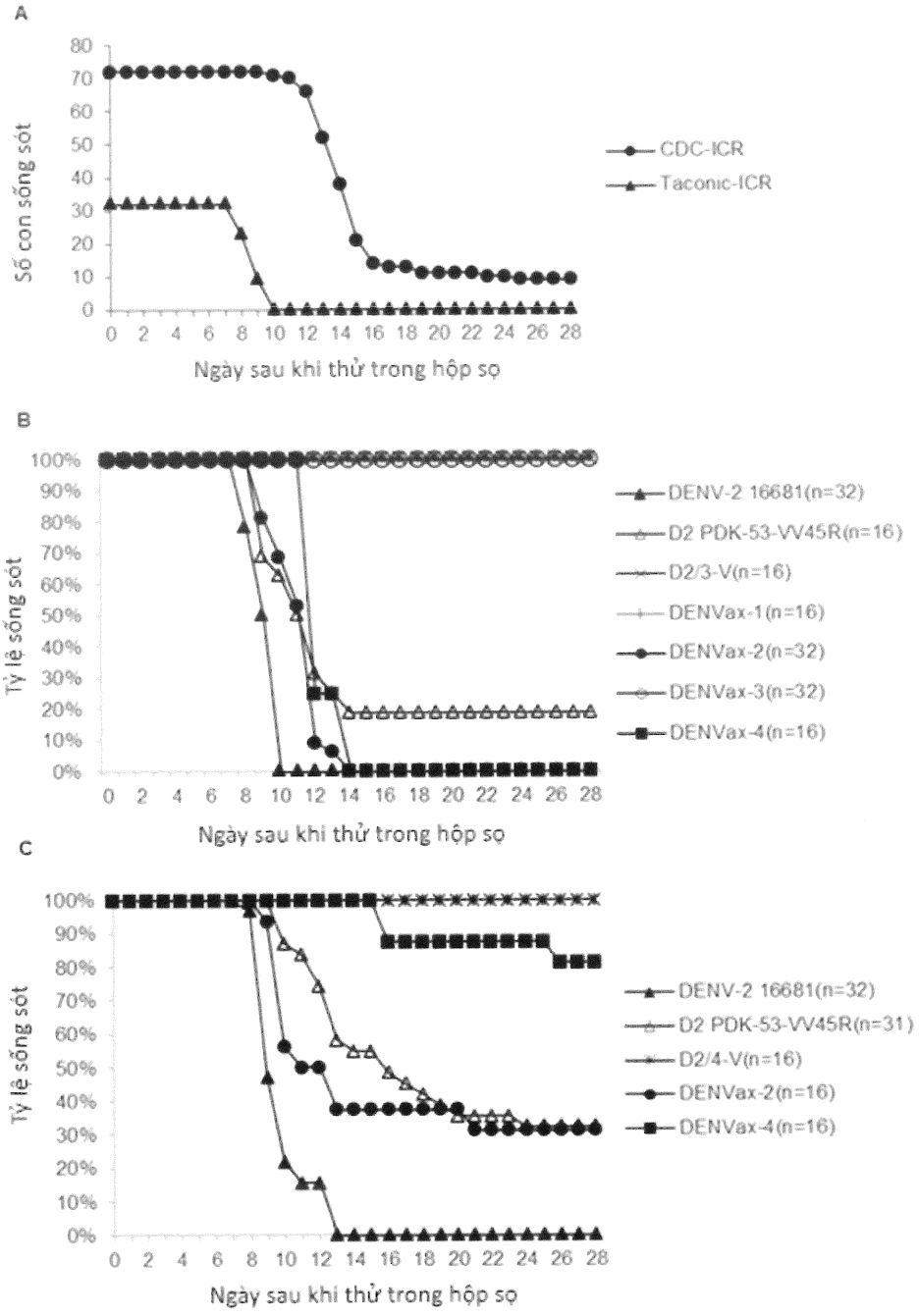

A
Số con sống sót
Ngày sau khi thử trong hộp sọ
CDC-ICR
Taconic-ICR
B
Tỷ lệ sống sót
Ngày sau khi thử trong hộp sọ
DENV-2 16681(n=32)
D2 PDK-53-VV45R(n=16)
D2/3-V(n=16)
DENVax-1(n=16)
DENVax-2(n=32)
DENVax-3(n=32)
DENVax-4(n=16)
C
Tỷ lệ sống sót
Ngày sau khi thử trong hộp sọ
DENV-2 16681(n=32)
D2 PDK-53-VV45R(n=31)
D2/4-V(n=16)
DENVax-2(n=16)
DENVax-4(n=16)

Fig. 6

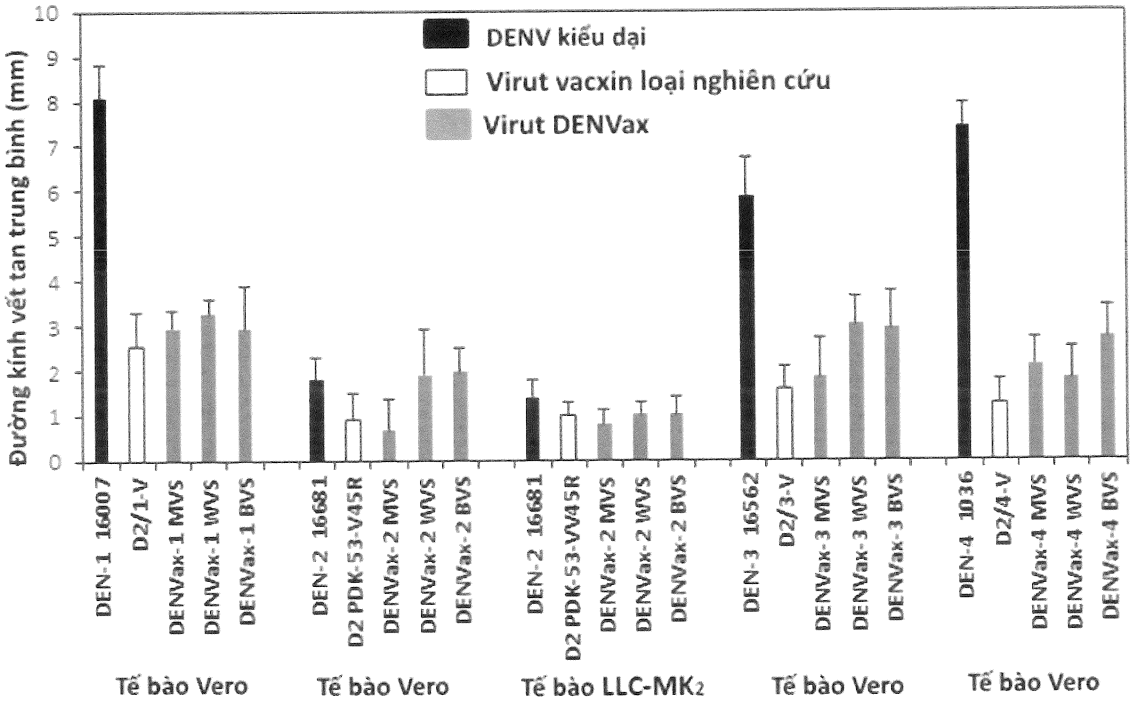

Đường kính vết tan trung bình (mm)
DENV kiểu dại
Virut vacxin loại nghiên cứu
Virut DENVax
DEN-1 16007
D2/1-V
DENVax-1 MVS
DENVax-1 WVS
DENVax-1 BVS
DEN-2 16681
D2 PDK-53-V45R
DENVax-2 MVS
DENVax-2 WVS
DENVax-2 BVS
DEN-2 16681
D2 PDK-53-VV45R
DENVax-2 MVS
DENVax-2 WVS
DENVax-2 BVS
DEN-3 16562
D2/3-V
DENVax-3 MVS
DENVax-3 WVS
DENVax-3 BVS
DEN-4 1036
D2/4-V
DENVax-4 MVS
DENVax-4 WVS
DENVax-4 BVS
Tế bào Vero
Tế bào Vero
Tế bào LLC-MK2
Tế bào Vero
Tế bào Vero

Fig. 7

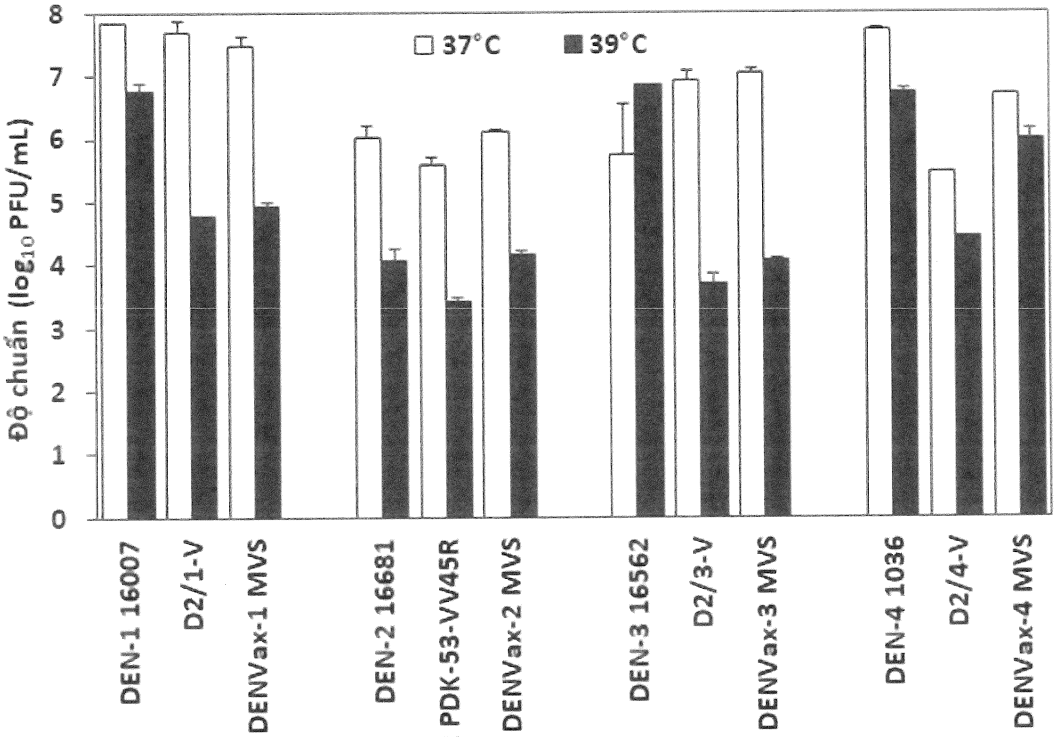

37°C
39°C
Độ chuẩn (log10 PFU/mL)
0
1
2
3
4
5
6
7
8
DEN-1 16007
D2/1-V
DENVax-1 MVS
DEN-2 16681
D2 PDK-53-VV45R
DENVax-2 MVS
DEN-3 16562
D2/3-V
DENVax-3 MVS
DEN-4 1036
D2/4-V
DENVax-4 MVS

Fig. 8

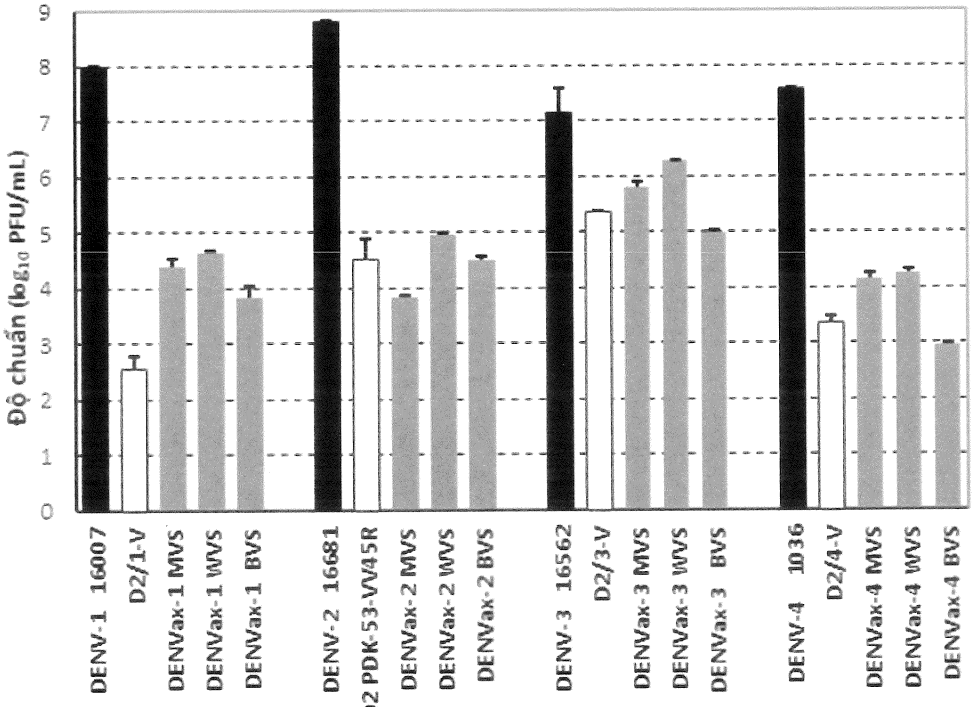

Độ chuẩn (log_{10} PFU/mL)
0
1
2
3
4
5
6
7
8
9
DENV-1 16007
D2/1-V
DENVax-1 MVS
DENVax-1 WVS
DENVax-1 BVS
DENV-2 16681
D2 PDK-53-VV45R
DENVax-2 MVS
DENVax-2 WVS
DENVax-2 BVS
DENV-3 16562
D2/3-V
DENVax-3 MVS
DENVax-3 WVS
DENVax-3 BVS
DENV-4 1036
D2/4-V
DENVax-4 MVS
DENVax-4 WVS
DENVax-4 BVS

Fig. 9A và 9B

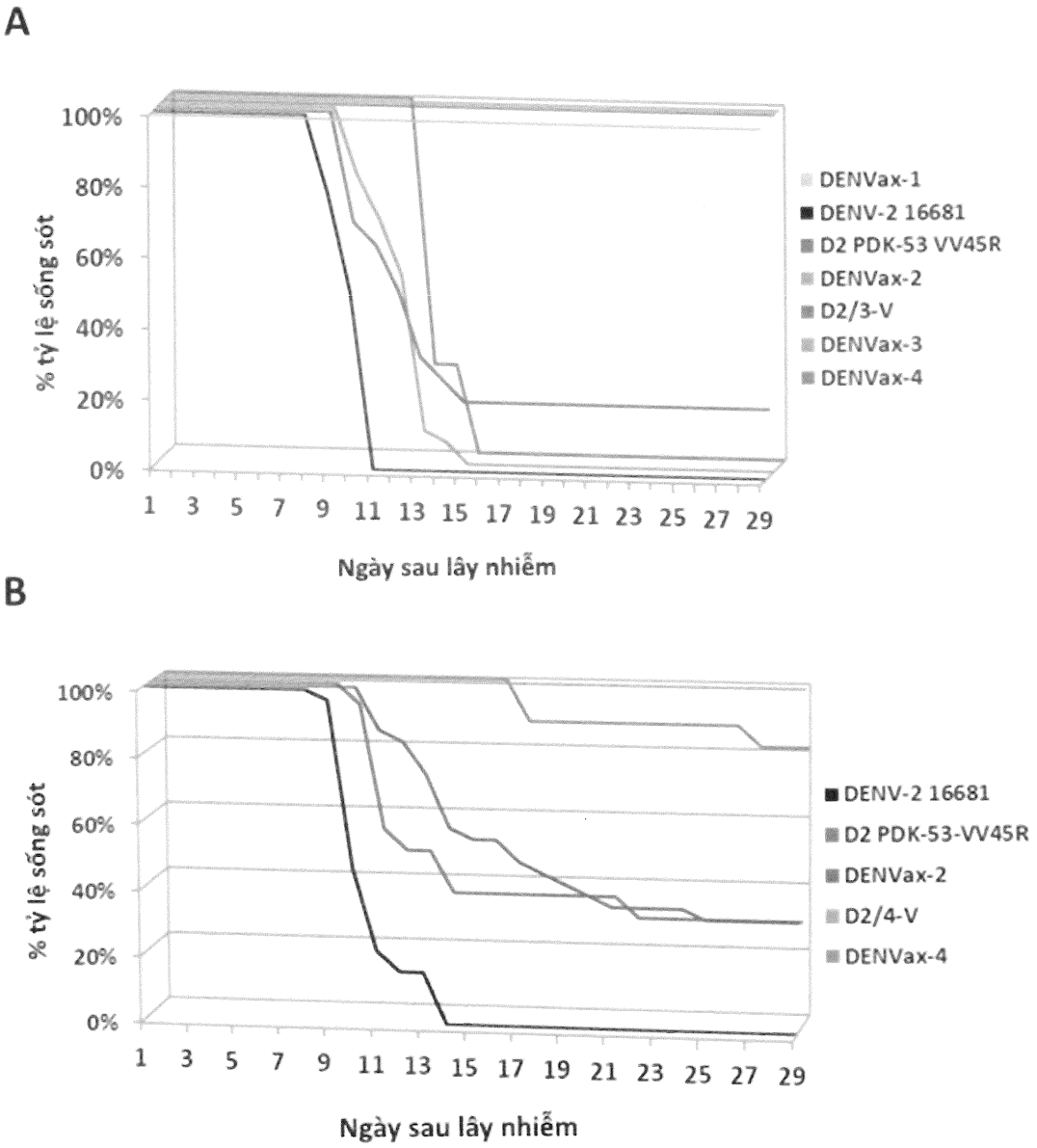

A
100%
80%
60%
40%
20%
0%
% tỷ lệ sống sót
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29
Ngày sau lây nhiễm
DENVax-1
DENV-2 16681
D2 PDK-53 VV45R
DENVax-2
D2/3-V
DENVax-3
DENVax-4
B
100%
80%
60%
40%
20%
0%
% tỷ lệ sống sót
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29
Ngày sau lây nhiễm
DENV-2 16681
D2 PDK-53-VV45R
DENVax-2
D2/4-V
DENVax-4

Fig. 10

Typ huyết thanh	Chủng	Nguồn virut	C57T 5'NCR	A524T prM-D29V	T900C[a] M (cảm)	C2055T E (cảm)	G2579A NS3-G53D	C4018T NS2A-L181	A5270 NS3-E350V	T5547C NS3 (cảm)	G6599C NS4A-G75A	C8571T NS5 (cảm)
DENV-2	16681	Phân lập từ người	C	A	T	C	G	C	A	T	G	C
	PDK-53	Cấy chuyền tế bào PDK của 1668	T	T	..	t	A	T	hỗn hợp T/A	c	C	t
	PDK53-V(VV45R)	PDK-53-V tái tổ hợp	T	T	c	t	A	T	T	c	C	C
	PDK53-E(VE48R)	PDK 53-E tái tổ hợp	T	T	c	t	A	T	A	c	C	C

Thế đột biến được gạch chân: 3 locus giảm độc lực quan trọng nhất của PDK-53
Font chữ màu đỏ: trình tự đặc hiệu PDK-53 (thay đổi từ 16681)
Font in đậm: Trình tự nt khác nhau giữa virut PDK-53 và dòng có nguồn gốc virut V hoặc E
[a] Gen đánh dấu dòng cảm được tạo ra để biệt hóa PDK-53 gốc và virut (dòng có nguồn gốc) tái tổ hợp

FIG. 11A

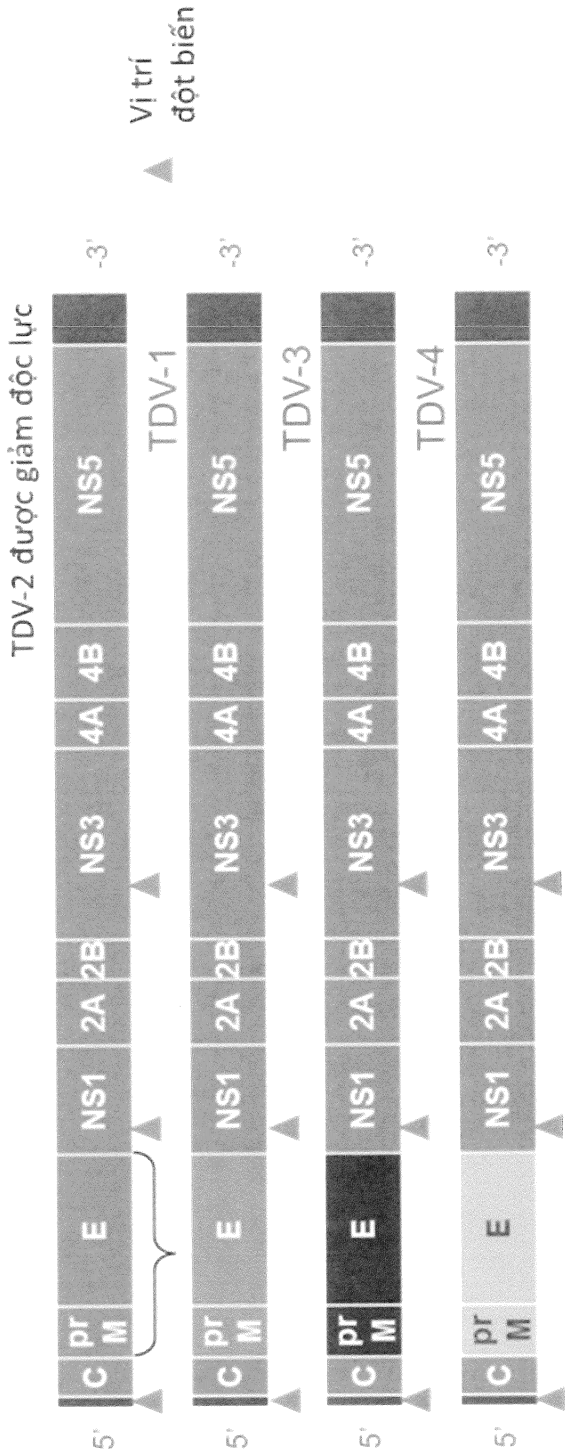

Vị trí đột biến
TDV-2 được giảm độc lực
TDV-1
TDV-3
TDV-4
5'
-3'
C
prM
E
NS1
2A
2B
NS3
4A
4B
NS5

FIG. 11B

Trên một liều	TDV-1	TDV-2	TDV-3	TDV-4	Tổng số
Chế phẩm TDV (Đơn vị hình thành vết tan)	2×10^4 PFU	5×10^4 PFU	1×10^5 PFU	3×10^5 PFU	$4,7 \times 10^5$ PFU

FIG. 12A

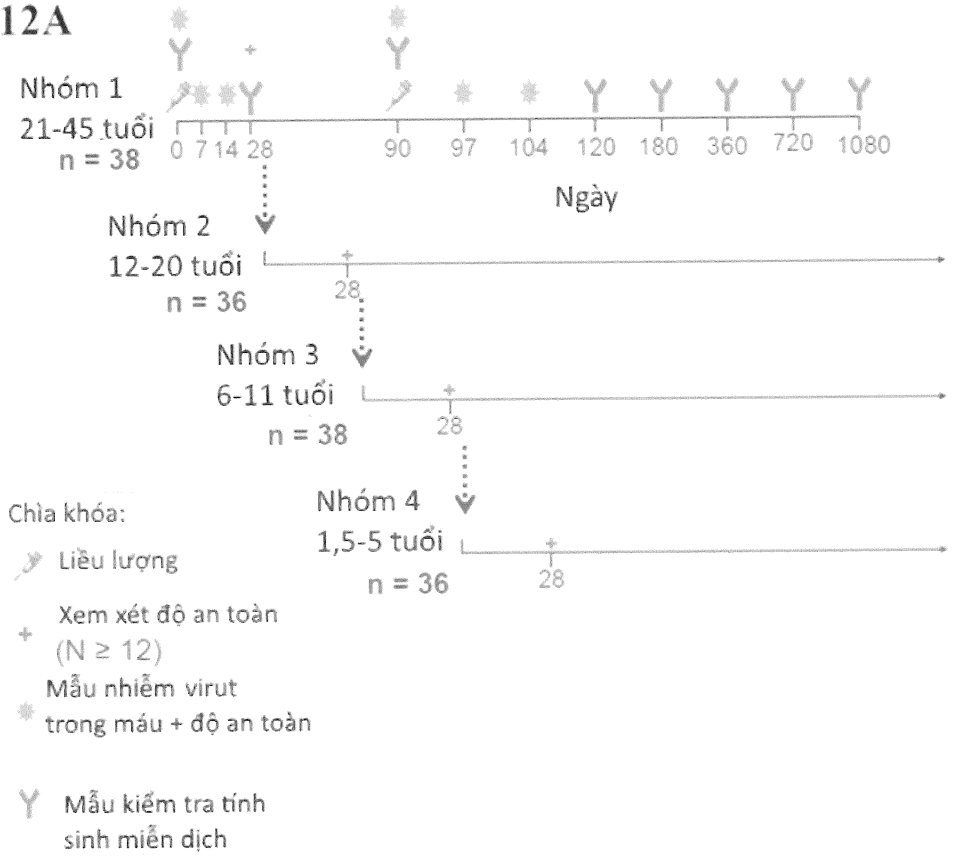

Nhóm 1
21-45 tuổi
n = 38
0 7 14 28 90 97 104 120 180 360 720 1080
Ngày
Nhóm 2
12-20 tuổi
n = 36
28
Nhóm 3
6-11 tuổi
n = 38
28
Nhóm 4
1,5-5 tuổi
n = 36
28
Chìa khóa:
Liều lượng
Xem xét độ an toàn
(N ≥ 12)
Mẫu nhiễm virut
trong máu + độ an toàn
Mẫu kiểm tra tính
sinh miễn dịch

FIG. 12B

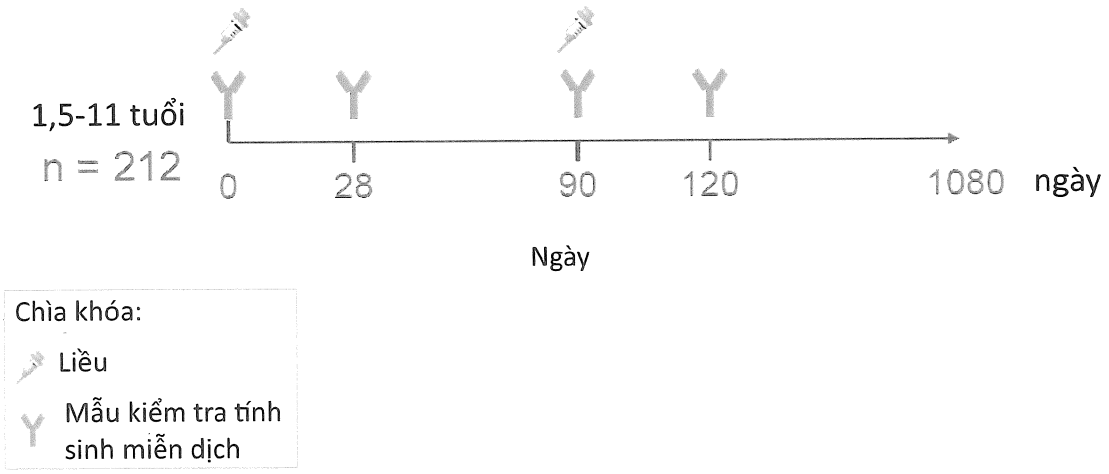

1,5-11 tuổi
n = 212
0
28
90
120
1080
ngày
Ngày
Chìa khóa:
Liều
Mẫu kiểm tra tính sinh miễn dịch

FIG. 13A

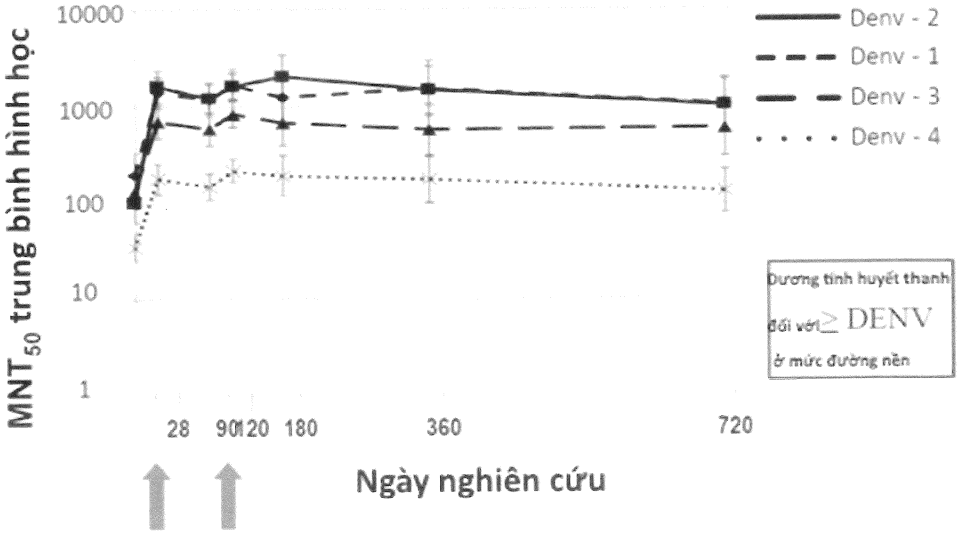

10000
1000
100
10
1
MNT50 trung bình hình học
28
90
120
180
360
720
Ngày nghiên cứu
Denv - 2
Denv - 1
Denv - 3
Denv - 4
Dương tính huyết thanh
đối với ≥ DENV
ở mức đường nền

FIG. 13B

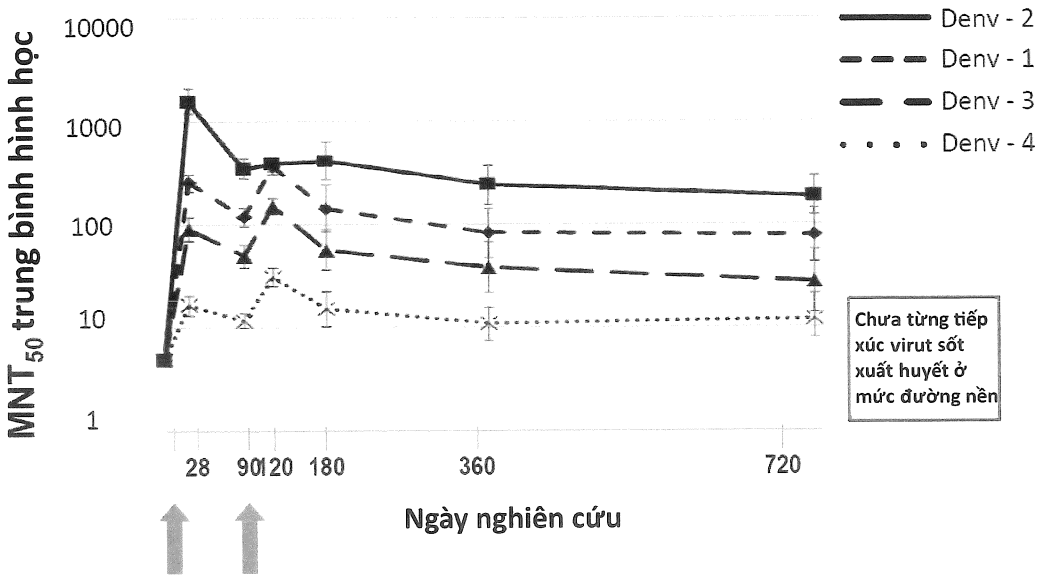

MNT50 trung bình hình học
10000
1000
100
10
1
28
90
120
180
360
720
Ngày nghiên cứu
Denv - 2
Denv - 1
Denv - 3
Denv - 4
Chưa từng tiếp xúc virut sốt xuất huyết ở mức đường nền

FIG. 14

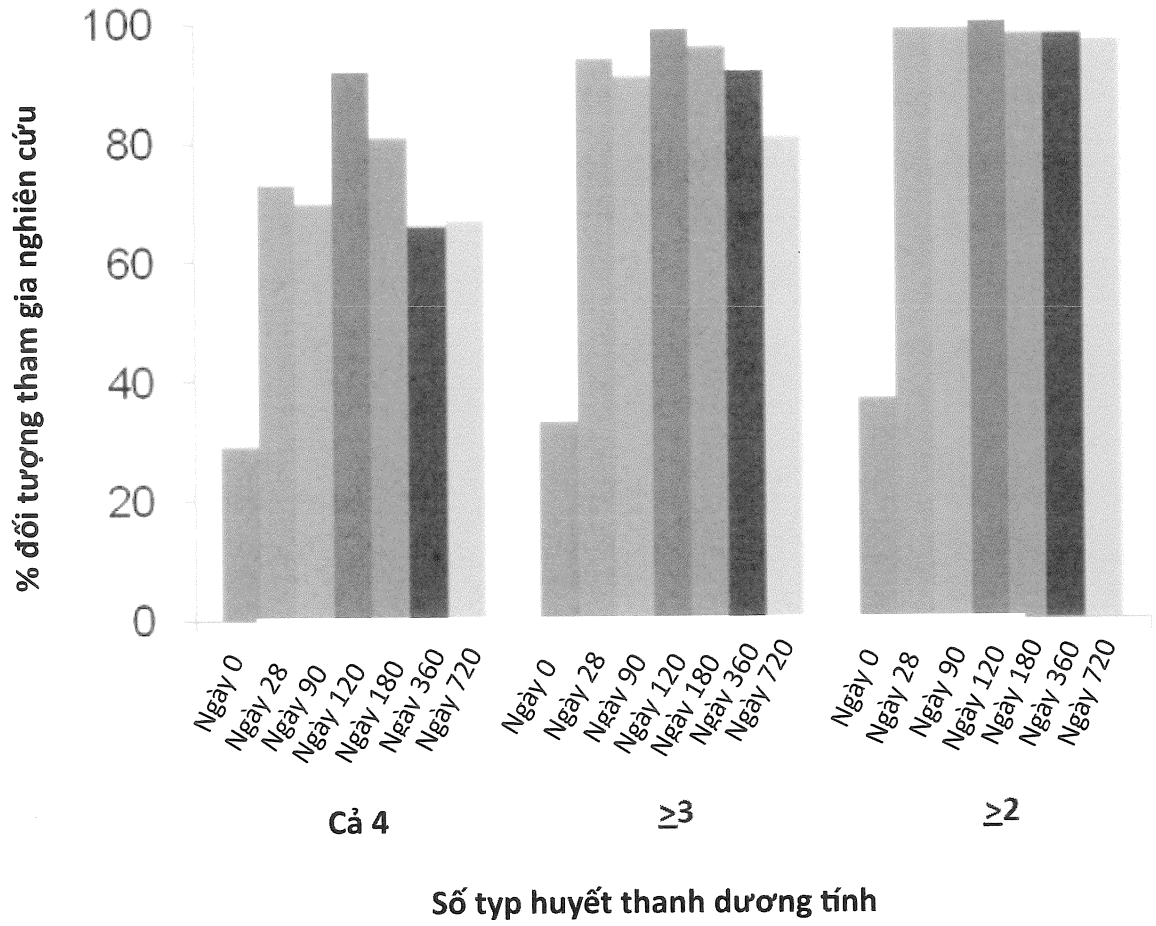

100
80
60
40
20
0
% đối tượng tham gia nghiên cứu
Ngày 0
Ngày 28
Ngày 90
Ngày 120
Ngày 180
Ngày 360
Ngày 720
Cả 4
≥3
≥2
Số typ huyết thanh dương tính

FIG. 15

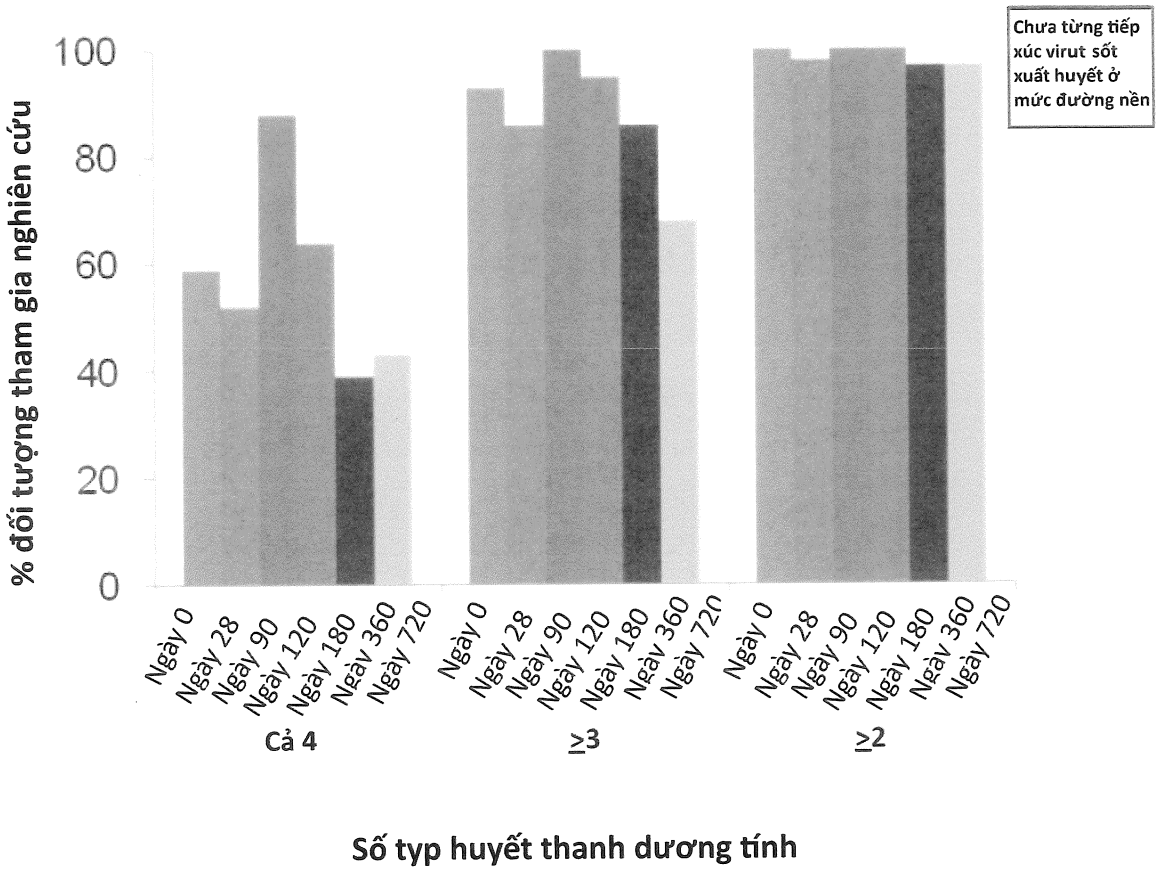
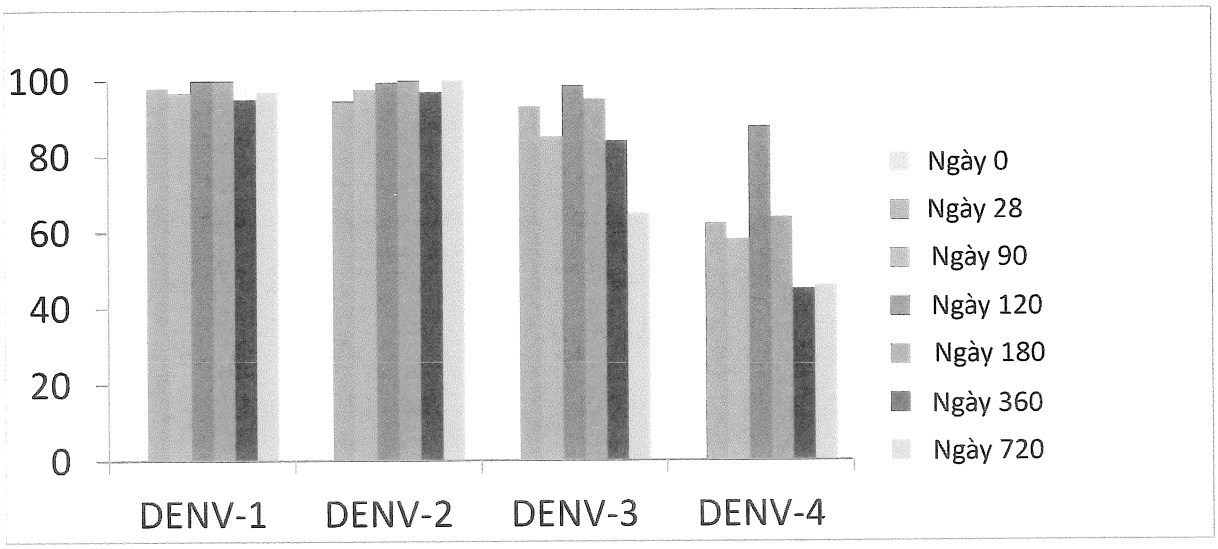
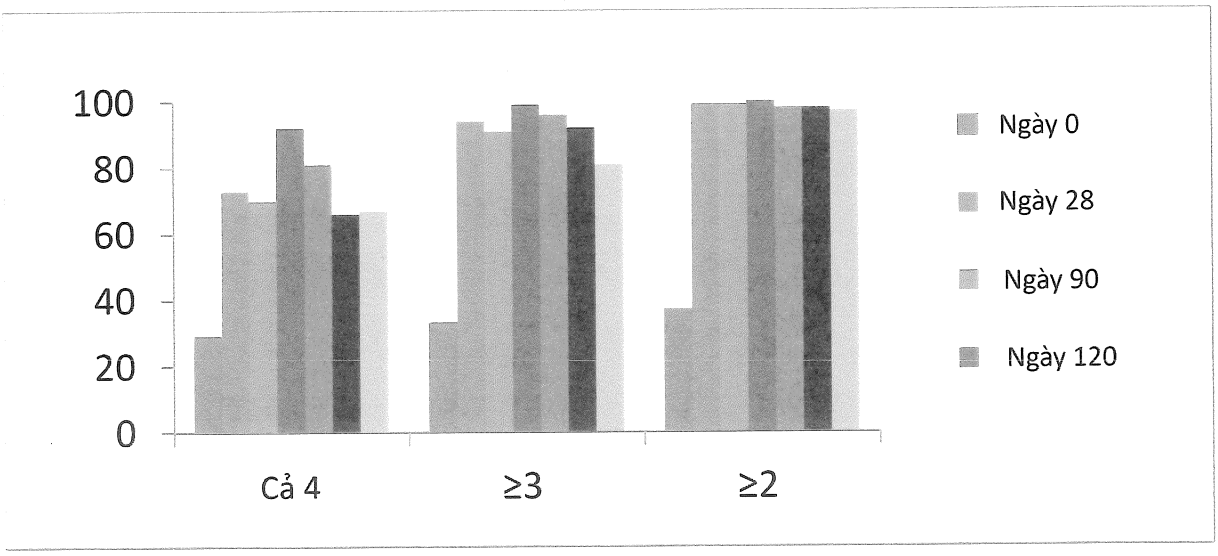

Chưa từng tiếp xúc virut sốt xuất huyết ở mức đường nền
% đối tượng tham gia nghiên cứu
100
80
60
40
20
0
Ngày 0
Ngày 28
Ngày 90
Ngày 120
Ngày 180
Ngày 360
Ngày 720
Cả 4
≥3
≥2
Số typ huyết thanh dương tính

FIG. 16



*Trẻ em/thanh niên chưa từng tiếp xúc virut sốt xuất huyết ở mức đường nền và sau khi nhận TDV (tứ giá)

FIG. 17



*Trẻ em/thanh niên chưa từng tiếp xúc virut ở mức đường nền và sau khi nhận TDV (tứ giá)

DANH MỤC TRÌNH TỰ

<110> Takeda Vaccines, Inc.
Wallace, Derek

<120> Dược phẩm chứa chế phẩm vacxin tứ giá để điều trị cho trẻ em hoặc thanh niên chống lại sự nhiễm virut sốt xuất huyết

<130> 478484.265

<150> XXXXXXXXX
<151> 2017-04-13

<160> 34

<170> PatentIn version 3.5

<210> 1
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, Tiền chủng gốc

<400> 1
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgggg gagagccgca tatgatagtt 480

agcaagcagg aaagaggaaa gtcacttttg ttcaagacct ctgcaggtgt caacatgtgc 540

accctcattg cgatggattt gggagagttg tgtgaggaca cgatgaccta caaatgcccc 600

cggatcactg aggcggaacc agatgacgtt gactgttggt gcaatgccac ggacacatgg 660

gtgacctatg gaacgtgctc tcaaactggc gaacaccgac gagacaaacg ttccgtcgca 720

ttggccccac acgtggggct tggcctagaa acaagagccg aaacgtggat gtcctctgaa 780

ggtgcttgga aacagataca aaaagtagag acttgggctc tgagacatcc aggattcacg 840

gtgatagccc tttttctagc acatgccata ggaacatcca tcacccagaa agggatcatt 900

ttcattttgc tgatgctggt aacaccatct atggccatgc gatgcgtggg aataggcaac 960

agagacttcg tggaaggact gtcaggagca acatgggtgg atgtggtact ggagcatgga 1020

agttgcgtca ccaccatggc aaaaaacaaa ccaacactgg acattgaact cttgaagacg 1080

gaggtcacaa accctgcagt tctgcgtaaa ttgtgcattg aagctaaaat atcaaacacc 1140

accaccgatt cgagatgtcc aacacaagga gaagccacac tggtggaaga acaagacgcg 1200

aactttgtgt gccgacgaac gttcgtggac agaggctggg gcaatggctg tgggctattc 1260

ggaaaaggta gtctaataac gtgtgccaag tttaagtgtg tgacaaaact agaaggaaag 1320

atagttcaat atgaaaacct aaaatattca gtgatagtca ccgtccacac tggagatcag 1380

caccaggtgg gaaatgagac tacagaacat ggaacaactg caaccataac acctcaagct 1440

cctacgtcgg aaatacagct gaccgactac ggaacccttа cattagattg ttcacctagg 1500

acagggctag attttaacga gatggtgttg ctgacaatga aagaaagatc atggcttgtc 1560

cacaaacaat ggttcctaga cttaccactg ccttggacct ctggggcttc aacatcccaa 1620

gagacttgga acagacaaga tttactggtc acatttaaga cagctcatgc aaagaagcag 1680

gaagtagtcg tactaggatc acaagaagga gcaatgcaca ctgcgctgac tggagcgaca 1740

gaaatccaaa cgtcaggaac gacaacaatt ttcgcaggac acctaaaatg cagactaaaa 1800

atggacaaac taactttaaa agggatgtca tatgtgatgt gcacaggctc attcaagtta 1860

gagaaagaag tggctgagac ccagcatgga actgttctgg tgcaggttaa atatgaagga 1920

acagacgcac catgcaagat tccctttcg acccaagatg agaaaggagc aacccagaat 1980

gggagattaa taacagccaa ccccatagtc actgacaaag aaaaaccagt caatattgag 2040

gcagaaccac cctttggtga gagctacatc gtggtaggag caggtgaaaa agctttgaaa 2100

ctaagctggt tcaagaaagg aagcagcata gggaaaatgt ttgaagcaac tgcccgagga 2160

gcacgaagga tggccattct gggagacacc gcatgggact tcggttctat aggaggagtg 2220

ttcacgtcta tgggaaaact ggtacaccag gtttttggaa ctgcatatgg agttttgttt 2280

agcggagttt cttggaccat gaaaatagga atagggattc tgctgacatg gctaggatta 2340

aattcaagga acacgtccct ttcgatgatg tgcatcgcag ccggcattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccctaccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagatcaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gtttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc gaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 2
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, Tiền chủng gốc

<400> 2

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Gly Gly Glu Pro His Met Ile Val
115 120 125

Ser Lys Gln Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly
130 135 140

Val Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Leu Cys Glu
145 150 155 160

Asp Thr Met Thr Tyr Lys Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp
165 170 175

Asp Val Asp Cys Trp Cys Asn Ala Thr Asp Thr Trp Val Thr Tyr Gly
180 185 190

Thr Cys Ser Gln Thr Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
195 200 205

Leu Ala Pro His Val Gly Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp
210 215 220

Met Ser Ser Glu Gly Ala Trp Lys Gln Ile Gln Lys Val Glu Thr Trp
225 230 235 240

Ala Leu Arg His Pro Gly Phe Thr Val Ile Ala Leu Phe Leu Ala His
245 250 255

Ala Ile Gly Thr Ser Ile Thr Gln Lys Gly Ile Ile Phe Ile Leu Leu
260 265 270

Met Leu Val Thr Pro Ser Met Ala Met Arg Cys Val Gly Ile Gly Asn
275 280 285

```
Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
    290                 295                 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Ile Glu Leu Leu Lys Thr Glu Val Thr Asn Pro Ala Val Leu
                325                 330                 335

Arg Lys Leu Cys Ile Glu Ala Lys Ile Ser Asn Thr Thr Thr Asp Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Ala Thr Leu Val Glu Glu Gln Asp Ala
        355                 360                 365

Asn Phe Val Cys Arg Arg Thr Phe Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Ile Thr Cys Ala Lys Phe Lys
385                 390                 395                 400

Cys Val Thr Lys Leu Glu Gly Lys Ile Val Gln Tyr Glu Asn Leu Lys
                405                 410                 415

Tyr Ser Val Ile Val Thr Val His Thr Gly Asp Gln His Gln Val Gly
            420                 425                 430

Asn Glu Thr Thr Glu His Gly Thr Thr Ala Thr Ile Thr Pro Gln Ala
        435                 440                 445

Pro Thr Ser Glu Ile Gln Leu Thr Asp Tyr Gly Thr Leu Thr Leu Asp
    450                 455                 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Thr
465                 470                 475                 480

Met Lys Glu Arg Ser Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
                485                 490                 495
```

```
Pro Leu Pro Trp Thr Ser Gly Ala Ser Thr Ser Gln Glu Thr Trp Asn
            500                 505                 510

Arg Gln Asp Leu Leu Val Thr Phe Lys Thr Ala His Ala Lys Lys Gln
        515                 520                 525

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

Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Thr Thr Thr Ile Phe Ala
545                 550                 555                 560

Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr Leu Lys Gly
                565                 570                 575

Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys Glu Val
            580                 585                 590

Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr Glu Gly
        595                 600                 605

Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu Lys Gly
    610                 615                 620

Ala Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val Thr Asp
625                 630                 635                 640

Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
                645                 650                 655

Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser Trp Phe
            660                 665                 670

Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly
        675                 680                 685

Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
    690                 695                 700
```

Ile Gly Gly Val Phe Thr Ser Met Gly Lys Leu Val His Gln Val Phe
705 710 715 720

Gly Thr Ala Tyr Gly Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
725 730 735

Ile Gly Ile Gly Ile Leu Leu Thr Trp Leu Gly Leu Asn Ser Arg Asn
740 745 750

Thr Ser Leu Ser Met Met Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
850 855 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865 870 875 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
885 890 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
915 920 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
930 935 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945 950 955 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
1025 1030 1035

Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp
1040 1045 1050

Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn
1055 1060 1065

Arg Gly Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile
1070 1075 1080

Thr Glu Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr
1085 1090 1095

Arg Gly Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu
1100 1105 1110

Lys Glu Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly
1115 1120 1125

His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala
1130 1135 1140

Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His
1145 1150 1155

Ala Ile Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly
1160 1165 1170

Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly
1175 1180 1185

Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala
1190 1195 1200

Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu
1205 1210 1215

Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr Ile
1220 1225 1230

Gly Ile Val Leu Leu Ser Gln Ser Thr Leu Pro Glu Thr Ile Leu
1235 1240 1245

Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met
1250 1255 1260

Val Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala
1265 1270 1275

Ile Leu Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys
1280 1285 1290

Val Ser Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe
1295 1300 1305

Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu
1310 1315 1320

Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu
1325 1330 1335

Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile
1340 1345 1350

Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys
1355 1360 1365

Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu
1370 1375 1380

Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu
1385 1390 1395

Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser
1400 1405 1410

Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser
1415 1420 1425

Met Ser Ile Lys Asn Glu Glu Glu Asp Gln Thr Leu Thr Ile Leu
1430 1435 1440

Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
1445 1450 1455

Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys
1460 1465 1470

Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met
1475 1480 1485

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys
1490 1495 1500

Gly Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu
1505 1510 1515

Gly Thr Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu
1520 1525 1530

Met His Lys Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys
1535 1540 1545

Lys Asp Leu Ile Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu
1550 1555 1560

Trp Lys Glu Gly Glu Glu Val Gln Val Leu Ala Leu Glu Pro Gly
1565 1570 1575

Lys Asn Pro Arg Ala Val Gln Thr Lys Pro Gly Leu Phe Lys Thr
1580 1585 1590

Asn Ala Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser Pro Gly
1595 1600 1605

Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly
1610 1615 1620

Leu Tyr Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser
1625 1630 1635

Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile
1640 1645 1650

Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu
1655 1660 1665

His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val
1670 1675 1680

Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro
1685 1690 1695

Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu
1700 1705 1710

Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly
1715 1720 1725

Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg
1730 1735 1740

Leu Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met
1745 1750 1755

Asp Glu Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly
1760 1765 1770

Tyr Ile Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe
1775 1780 1785

Met Thr Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser
1790 1795 1800

Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser
1805 1810 1815

Trp Asn Ser Gly His Glu Trp Val Thr Asp Phe Lys Gly Lys Thr
1820 1825 1830

Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala
1835 1840 1845

Cys Leu Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg Lys
1850 1855 1860

Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp
1865 1870 1875

Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys
1880 1885 1890

Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile
1895 1900 1905

Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro
1910 1915 1920

Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
1925 1930 1935

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
1955 1960 1965

Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met
1970 1975 1980

Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr
1985 1990 1995

Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg
2000 2005 2010

Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu
2015 2020 2025

Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys
2030 2035 2040

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

```
Ala Ala  Gly Arg Lys Ser Leu  Thr Leu Asn Leu Ile  Thr Glu Met
    2090                 2095                 2100

Gly Arg  Leu Pro Thr Phe Met  Thr Gln Lys Ala Arg  Asp Ala Leu
    2105                 2110                 2115

Asp Asn  Leu Ala Val Leu His  Thr Ala Glu Ala Gly  Gly Arg Ala
    2120                 2125                 2130

Tyr Asn  His Ala Leu Ser Glu  Leu Pro Glu Thr Leu  Glu Thr Leu
    2135                 2140                 2145

Leu Leu  Leu Thr Leu Leu Ala  Thr Val Thr Gly Gly  Ile Phe Leu
    2150                 2155                 2160

Phe Leu  Met Ser Ala Arg Gly  Ile Gly Lys Met Thr  Leu Gly Met
    2165                 2170                 2175

Cys Cys  Ile Ile Thr Ala Ser  Ile Leu Leu Trp Tyr  Ala Gln Ile
    2180                 2185                 2190

Gln Pro  His Trp Ile Ala Ala  Ser Ile Ile Leu Glu  Phe Phe Leu
    2195                 2200                 2205

Ile Val  Leu Leu Ile Pro Glu  Pro Glu Lys Gln Arg  Thr Pro Gln
    2210                 2215                 2220

Asp Asn  Gln Leu Thr Tyr Val  Val Ile Ala Ile Leu  Thr Val Val
    2225                 2230                 2235

Ala Ala  Thr Met Ala Asn Glu  Met Gly Phe Leu Glu  Lys Thr Lys
    2240                 2245                 2250

Lys Asp  Leu Gly Leu Gly Ser  Ile Ala Thr Gln Gln  Pro Glu Ser
    2255                 2260                 2265

Asn Ile  Leu Asp Ile Asp Leu  Arg Pro Ala Ser Ala  Trp Thr Leu
    2270                 2275                 2280
```

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg
2525 2530 2535

Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu
2540 2545 2550

Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val
2555 2560 2565

Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly
2570 2575 2580

Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly
2585 2590 2595

Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn
2600 2605 2610

Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro
2615 2620 2625

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp
2810 2815 2820

Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr
2825 2830 2835

Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
2840 2845 2850

Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr
2855 2860 2865

Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg
2870 2875 2880

Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
2885 2890 2895

Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala
2900 2905 2910

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys
2945 2950 2955

Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg
2960 2965 2970

Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp
2975 2980 2985

Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu
2990 2995 3000

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
3080 3085 3090

Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala
3095 3100 3105

Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile
3110 3115 3120

Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
3125 3130 3135

Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly
3140 3145 3150

Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
3155 3160 3165

Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln
3170 3175 3180

Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro
3185 3190 3195

Phe Cys Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg
3200 3205 3210

Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg
3215 3220 3225

Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala
3230 3235 3240

Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe
3245 3250 3255

His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
3260 3265 3270

Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
3275 3280 3285

His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
3290 3295 3300

Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 3
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, MVS

<400> 3
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgggg gagagccgca tatgatagtt 480

agcaagcagg aaagaggaaa gtcacttttg ttcaagacct ctgcaggtgt caacatgtgc 540

accctcattg cgatggattt gggagagttg tgtgaggaca cgatgaccta caaatgcccc 600

cggatcactg aggcggaacc agatgacgtt gactgttggt gcaatgccac ggacacatgg 660

gtgacctatg gaacgtgctc tcaaactggc gaacaccgac gagacaaacg ttccgtcgca 720

ttggccccac acgtggggct tggcctagaa acaagagccg aaacgtggat gtcctctgaa 780

ggtgcttgga aacagataca aaaagtagag acttgggctc tgagacatcc aggattcacg 840

gtgatagccc tttttctagc acatgccata ggaacatcca tcacccagaa agggatcatt 900

ttcattttgc tgatgctggt aacaccatct atggccatgc gatgcgtggg aataggcaac 960

agagacttcg tggaaggact gtcaggagca acatgggtgg atgtggtact ggagcatgga 1020

agttgcgtca ccaccatggc aaaaaacaaa ccaacactgg acattgaact cttgaagacg 1080

gaggtcacaa accctgcagt tctgcgtaaa ttgtgcattg aagctaaaat atcaaacacc 1140

accaccgatt cgagatgtcc aacacaagga gaagccacac tggtggaaga acaagacgcg 1200

aactttgtgt gccgacgaac gttcgtggac agaggctggg gcaatggctg tgggctattc 1260

ggaaaaggta gtctaataac gtgtgccaag tttaagtgtg tgacaaaact agaaggaaag 1320

atagttcaat atgaaaacct aaaatattca gtgatagtca ccgtccacac tggagatcag 1380

caccaggtgg gaaatgagac tacagaacat ggaacaactg caaccataac acctcaagct 1440

cctacgtcgg aaatacagct gaccgactac ggaaccctta cattagattg ttcacctagg 1500

acagggctag attttaacga gatggtgttg ctgacaatga aagaaagatc atggcttgtc 1560

cacaaacaat ggttcctaga cttaccactg ccttggacct ctggggcttc aacatcccaa 1620

gagacttgga acagacaaga tttactggtc acatttaaga cagctcatgc aaagaagcag 1680

gaagtagtcg tactaggatc acaagaagga gcaatgcaca ctgcgctgac tggagcgaca 1740

gaaatccaaa cgtcaggaac gacaacaatt ttcgcaggac acctaaaatg cagactaaaa 1800

atggacaaac taactttaaa agggatgtca tatgtgatgt gcacaggctc attcaagtta 1860

gagaaagaag tggctgagac ccagcatgga actgttctgg tgcaggttaa atatgaagga 1920

acagacgcac catgcaagat tccctttcg acccaagatg agaaaggagc aacccagaat 1980

gggagattaa taacagccaa ccccatagtc actgacaaag aaaaaccagt caatattgag 2040

gcagaaccac cctttggtga gagctacatc gtggtaggag caggtgaaaa agctttgaaa 2100

ctaagctggt tcaagaaagg aagcagcata gggaaaatgt ttgaagcaac tgcccgagga 2160

gcacgaagga tggccattct gggagacacc gcatgggact tcggttctat aggaggagtg 2220

ttcacgtcta tgggaaaact ggtacaccag gtttttggaa ctgcatatgg agttttgttt 2280

agcggagttt cttggaccat gaaaatagga atagggattc tgctgacatg gctaggatta 2340

aattcaagga acacgtccct ttcgatgatg tgcatcgcag ccggcattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccctaccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagatcaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gtttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc gaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 4
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, MVS

<400> 4

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

```
Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
            20                  25                  30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
        35                  40                  45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
    50                  55                  60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65                  70                  75                  80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
                85                  90                  95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
            100                 105                 110

Met Ala Phe His Leu Thr Thr Arg Gly Gly Glu Pro His Met Ile Val
        115                 120                 125

Ser Lys Gln Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly
    130                 135                 140

Val Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Leu Cys Glu
145                 150                 155                 160

Asp Thr Met Thr Tyr Lys Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp
                165                 170                 175

Asp Val Asp Cys Trp Cys Asn Ala Thr Asp Thr Trp Val Thr Tyr Gly
            180                 185                 190

Thr Cys Ser Gln Thr Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
        195                 200                 205

Leu Ala Pro His Val Gly Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp
    210                 215                 220
```

Met Ser Ser Glu Gly Ala Trp Lys Gln Ile Gln Lys Val Glu Thr Trp
225 230 235 240

Ala Leu Arg His Pro Gly Phe Thr Val Ile Ala Leu Phe Leu Ala His
245 250 255

Ala Ile Gly Thr Ser Ile Thr Gln Lys Gly Ile Ile Phe Ile Leu Leu
260 265 270

Met Leu Val Thr Pro Ser Met Ala Met Arg Cys Val Gly Ile Gly Asn
275 280 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
290 295 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305 310 315 320

Leu Asp Ile Glu Leu Leu Lys Thr Glu Val Thr Asn Pro Ala Val Leu
325 330 335

Arg Lys Leu Cys Ile Glu Ala Lys Ile Ser Asn Thr Thr Thr Asp Ser
340 345 350

Arg Cys Pro Thr Gln Gly Glu Ala Thr Leu Val Glu Glu Gln Asp Ala
355 360 365

Asn Phe Val Cys Arg Arg Thr Phe Val Asp Arg Gly Trp Gly Asn Gly
370 375 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Ile Thr Cys Ala Lys Phe Lys
385 390 395 400

Cys Val Thr Lys Leu Glu Gly Lys Ile Val Gln Tyr Glu Asn Leu Lys
405 410 415

Tyr Ser Val Ile Val Thr Val His Thr Gly Asp Gln His Gln Val Gly
420 425 430

```
Asn Glu Thr Thr Glu His Gly Thr Thr Ala Thr Ile Thr Pro Gln Ala
        435                 440                 445

Pro Thr Ser Glu Ile Gln Leu Thr Asp Tyr Gly Thr Leu Thr Leu Asp
    450                 455                 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Thr
465                 470                 475                 480

Met Lys Glu Arg Ser Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
                485                 490                 495

Pro Leu Pro Trp Thr Ser Gly Ala Ser Thr Ser Gln Glu Thr Trp Asn
            500                 505                 510

Arg Gln Asp Leu Leu Val Thr Phe Lys Thr Ala His Ala Lys Lys Gln
        515                 520                 525

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

Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Thr Thr Thr Ile Phe Ala
545                 550                 555                 560

Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr Leu Lys Gly
                565                 570                 575

Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys Glu Val
            580                 585                 590

Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr Glu Gly
        595                 600                 605

Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu Lys Gly
    610                 615                 620

Ala Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val Thr Asp
625                 630                 635                 640
```

Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
645 650 655

Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser Trp Phe
660 665 670

Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly
675 680 685

Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
690 695 700

Ile Gly Gly Val Phe Thr Ser Met Gly Lys Leu Val His Gln Val Phe
705 710 715 720

Gly Thr Ala Tyr Gly Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
725 730 735

Ile Gly Ile Gly Ile Leu Leu Thr Trp Leu Gly Leu Asn Ser Arg Asn
740 745 750

Thr Ser Leu Ser Met Met Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

```
Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
    850                 855                 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865                 870                 875                 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
                885                 890                 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
            900                 905                 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
        915                 920                 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
            980                 985                 990

Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035

Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050
```

```
Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065

Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                1075                1080

Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                1090                1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                1105                1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                1120                1125

His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                1135                1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                1150                1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                1165                1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                1180                1185

Ala Thr  Met Thr Asp Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                1195                1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                1210                1215

Leu Leu  Arg Lys Leu Thr Ser  Lys Glu Leu Met Met  Thr Thr Ile
    1220                1225                1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Leu Pro Glu  Thr Ile Leu
    1235                1240                1245
```

```
Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                 1255                 1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                 1270                 1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                 1285                 1290

Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                 1300                 1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                 1315                 1320

Thr Ile  Lys Gly Leu Asn Pro  Thr Ala Ile Phe Leu  Thr Thr Leu
    1325                 1330                 1335

Ser Arg  Thr Ser Lys Lys Arg  Ser Trp Pro Leu Asn  Glu Ala Ile
    1340                 1345                 1350

Met Ala  Val Gly Met Val Ser  Ile Leu Ala Ser Ser  Leu Leu Lys
    1355                 1360                 1365

Asn Asp  Ile Pro Met Thr Gly  Pro Leu Val Ala Gly  Gly Leu Leu
    1370                 1375                 1380

Thr Val  Cys Tyr Val Leu Thr  Gly Arg Ser Ala Asp  Leu Glu Leu
    1385                 1390                 1395

Glu Arg  Ala Ala Asp Val Lys  Trp Glu Asp Gln Ala  Glu Ile Ser
    1400                 1405                 1410

Gly Ser  Ser Pro Ile Leu Ser  Ile Thr Ile Ser Glu  Asp Gly Ser
    1415                 1420                 1425

Met Ser  Ile Lys Asn Glu Glu  Glu Asp Gln Thr Leu  Thr Ile Leu
    1430                 1435                 1440
```

```
Ile Arg  Thr Gly Leu Leu Val  Ile Ser Gly Leu Phe  Pro Val Ser
    1445                1450                1455

Ile Pro  Ile Thr Ala Ala Ala  Trp Tyr Leu Trp Glu  Val Lys Lys
    1460                1465                1470

Gln Arg  Ala Gly Val Leu Trp  Asp Val Pro Ser Pro  Pro Pro Met
    1475                1480                1485

Gly Lys  Ala Glu Leu Glu Asp  Gly Ala Tyr Arg Ile  Lys Gln Lys
    1490                1495                1500

Gly Ile  Leu Gly Tyr Ser Gln  Ile Gly Ala Gly Val  Tyr Lys Glu
    1505                1510                1515

Gly Thr  Phe His Thr Met Trp  His Val Thr Arg Gly  Ala Val Leu
    1520                1525                1530

Met His  Lys Gly Lys Arg Ile  Glu Pro Ser Trp Ala  Asp Val Lys
    1535                1540                1545

Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                1555                1560

Trp Lys  Glu Gly Glu Glu Val  Gln Val Leu Ala Leu  Glu Pro Gly
    1565                1570                1575

Lys Asn  Pro Arg Ala Val Gln  Thr Lys Pro Gly Leu  Phe Lys Thr
    1580                1585                1590

Asn Ala  Gly Thr Ile Gly Ala  Val Ser Leu Asp Phe  Ser Pro Gly
    1595                1600                1605

Thr Ser  Gly Ser Pro Ile Ile  Asp Lys Lys Gly Lys  Val Val Gly
    1610                1615                1620

Leu Tyr  Gly Asn Gly Val Val  Thr Arg Ser Gly Ala  Tyr Val Ser
    1625                1630                1635
```

```
Ala Ile  Ala Gln Thr Glu Lys  Ser Ile Glu Asp Asn  Pro Glu Ile
    1640                1645                1650

Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                1660                1665

His Pro  Gly Ala Gly Lys Thr  Lys Arg Tyr Leu Pro  Ala Ile Val
    1670                1675                1680

Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                1690                1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                1705                1710

Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                1720                1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                1735                1740

Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                1750                1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                1765                1770

Tyr Ile  Ser Thr Arg Val Glu  Met Gly Glu Ala Ala  Gly Ile Phe
    1775                1780                1785

Met Thr  Ala Thr Pro Pro Gly  Ser Arg Asp Pro Phe  Pro Gln Ser
    1790                1795                1800

Asn Ala  Pro Ile Ile Asp Glu  Glu Arg Glu Ile Pro  Glu Arg Ser
    1805                1810                1815

Trp Asn  Ser Gly His Glu Trp  Val Thr Asp Phe Lys  Gly Lys Thr
    1820                1825                1830
```

Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala
1835 1840 1845

Cys Leu Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg Lys
1850 1855 1860

Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp
1865 1870 1875

Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys
1880 1885 1890

Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile
1895 1900 1905

Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro
1910 1915 1920

Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
1925 1930 1935

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
1955 1960 1965

Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met
1970 1975 1980

Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr
1985 1990 1995

Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg
2000 2005 2010

Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu
2015 2020 2025

Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys
2030 2035 2040

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met
2090 2095 2100

Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu
2105 2110 2115

Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala
2120 2125 2130

Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu
2135 2140 2145

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

```
Leu Cys  Val Thr Gln Val Leu  Met Met Arg Thr Thr  Trp Ala Leu
    2420                 2425                 2430

Cys Glu  Ala Leu Thr Leu Ala  Thr Gly Pro Ile Ser  Thr Leu Trp
    2435                 2440                 2445

Glu Gly  Asn Pro Gly Arg Phe  Trp Asn Thr Thr Ile  Ala Val Ser
    2450                 2455                 2460

Met Ala  Asn Ile Phe Arg Gly  Ser Tyr Leu Ala Gly  Ala Gly Leu
    2465                 2470                 2475

Leu Phe  Ser Ile Met Lys Asn  Thr Thr Asn Thr Arg  Arg Gly Thr
    2480                 2485                 2490

Gly Asn  Ile Gly Glu Thr Leu  Gly Glu Lys Trp Lys  Ser Arg Leu
    2495                 2500                 2505

Asn Ala  Leu Gly Lys Ser Glu  Phe Gln Ile Tyr Lys  Lys Ser Gly
    2510                 2515                 2520

Ile Gln  Glu Val Asp Arg Thr  Leu Ala Lys Glu Gly  Ile Lys Arg
    2525                 2530                 2535

Gly Glu  Thr Asp His His Ala  Val Ser Arg Gly Ser  Ala Lys Leu
    2540                 2545                 2550

Arg Trp  Phe Val Glu Arg Asn  Met Val Thr Pro Glu  Gly Lys Val
    2555                 2560                 2565

Val Asp  Leu Gly Cys Gly Arg  Gly Gly Trp Ser Tyr  Tyr Cys Gly
    2570                 2575                 2580

Gly Leu  Lys Asn Val Arg Glu  Val Lys Gly Leu Thr  Lys Gly Gly
    2585                 2590                 2595

Pro Gly  His Glu Glu Pro Ile  Pro Met Ser Thr Tyr  Gly Trp Asn
    2600                 2605                 2610
```

Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro
2615 2620 2625

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

```
Ser Ser  Met Val Asn Gly Val  Val Arg Leu Leu Thr  Lys Pro Trp
    2810                 2815                 2820

Asp Val  Val Pro Met Val Thr  Gln Met Ala Met Thr  Asp Thr Thr
    2825                 2830                 2835

Pro Phe  Gly Gln Gln Arg Val  Phe Lys Glu Lys Val  Asp Thr Arg
    2840                 2845                 2850

Thr Gln  Glu Pro Lys Glu Gly  Thr Lys Lys Leu Met  Lys Ile Thr
    2855                 2860                 2865

Ala Glu  Trp Leu Trp Lys Glu  Leu Gly Lys Lys Lys  Thr Pro Arg
    2870                 2875                 2880

Met Cys  Thr Arg Glu Glu Phe  Thr Arg Lys Val Arg  Ser Asn Ala
    2885                 2890                 2895

Ala Leu  Gly Ala Ile Phe Thr  Asp Glu Asn Lys Trp  Lys Ser Ala
    2900                 2905                 2910

Arg Glu  Ala Val Glu Asp Ser  Arg Phe Trp Glu Leu  Val Asp Lys
    2915                 2920                 2925

Glu Arg  Asn Leu His Leu Glu  Gly Lys Cys Glu Thr  Cys Val Tyr
    2930                 2935                 2940

Asn Met  Met Gly Lys Arg Glu  Lys Lys Leu Gly Glu  Phe Gly Lys
    2945                 2950                 2955

Ala Lys  Gly Ser Arg Ala Ile  Trp Tyr Met Trp Leu  Gly Ala Arg
    2960                 2965                 2970

Phe Leu  Glu Phe Glu Ala Leu  Gly Phe Leu Asn Glu  Asp His Trp
    2975                 2980                 2985

Phe Ser  Arg Glu Asn Ser Leu  Ser Gly Val Glu Gly  Glu Gly Leu
    2990                 2995                 3000
```

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
3080 3085 3090

Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala
3095 3100 3105

Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile
3110 3115 3120

Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
3125 3130 3135

Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly
3140 3145 3150

Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
3155 3160 3165

Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln
3170 3175 3180

Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro
3185 3190 3195

Phe Cys Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg
3200 3205 3210

Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg
3215 3220 3225

Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala
3230 3235 3240

Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe
3245 3250 3255

His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
3260 3265 3270

Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
3275 3280 3285

His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
3290 3295 3300

Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 5
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, WVS

<220>
<221> dấu hiệu hỗn tạp
<222> (2384)..(2384)
<223> n là xytosin hoặc guanin

<400> 5
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgggg gagagccgca tatgatagtt 480

agcaagcagg aaagaggaaa gtcacttttg ttcaagacct ctgcaggtgt caacatgtgc 540

accctcattg cgatggattt gggagagttg tgtgaggaca cgatgaccta caaatgcccc 600

cggatcactg aggcggaacc agatgacgtt gactgttggt gcaatgccac ggacacatgg 660

gtgacctatg gaacgtgctc tcaaactggc gaacaccgac gagacaaacg ttccgtcgca 720

ttggccccac acgtggggct tggcctagaa acaagagccg aaacgtggat gtcctctgaa 780

ggtgcttgga aacagataca aaaagtagag acttgggctc tgagacatcc aggattcacg 840

gtgatagccc tttttctagc acatgccata ggaacatcca tcacccagaa agggatcatt 900

ttcattttgc tgatgctggt aacaccatct atggccatgc gatgcgtggg aataggcaac 960

agagacttcg tggaaggact gtcaggagca acatgggtgg atgtggtact ggagcatgga 1020

agttgcgtca ccaccatggc aaaaaacaaa ccaacactgg acattgaact cttgaagacg 1080

gaggtcacaa accctgcagt tctgcgtaaa ttgtgcattg aagctaaaat atcaaacacc 1140

accaccgatt cgagatgtcc aacacaagga gaagccacac tggtggaaga acaagacgcg 1200

aactttgtgt gccgacgaac gttcgtggac agaggctggg gcaatggctg tgggctattc 1260

ggaaaaggta gtctaataac gtgtgccaag tttaagtgtg tgacaaaact agaaggaaag 1320

atagttcaat atgaaaacct aaaatattca gtgatagtca ccgtccacac tggagatcag 1380

caccaggtgg gaaatgagac tacagaacat ggaacaactg caaccataac acctcaagct 1440

cctacgtcgg aaatacagct gaccgactac ggaaccctta cattagattg ttcacctagg 1500

acagggctag attttaacga gatggtgttg ctgacaatga aagaaagatc atggcttgtc 1560

cacaaacaat ggttcctaga cttaccactg ccttggacct ctggggcttc aacatcccaa 1620

gagacttgga acagacaaga tttactggtc acatttaaga cagctcatgc aaagaagcag 1680

gaagtagtcg tactaggatc acaagaagga gcaatgcaca ctgcgctgac tggagcgaca 1740

gaaatccaaa cgtcaggaac gacaacaatt ttcgcaggac acctaaaatg cagactaaaa 1800

atggacaaac taactttaaa agggatgtca tatgtgatgt gcacaggctc attcaagtta 1860

gagaaagaag tggctgagac ccagcatgga actgttctgg tgcaggttaa atatgaagga 1920

acagacgcac catgcaagat tccctttttcg acccaagatg agaaaggagc aacccagaat 1980

gggagattaa taacagccaa ccccatagtc actgacaaag aaaaaccagt caatattgag 2040

gcagaaccac cctttggtga gagctacatc gtggtaggag caggtgaaaa agctttgaaa 2100

ctaagctggt tcaagaaagg aagcagcata gggaaaatgt ttgaagcaac tgcccgagga 2160

gcacgaagga tggccattct gggagacacc gcatgggact tcggttctat aggaggagtg 2220

ttcacgtcta tgggaaaact ggtacaccag gtttttggaa ctgcatatgg agttttgttt 2280

agcggagttt cttggaccat gaaaatagga atagggattc tgctgacatg gctaggatta 2340

aattcaagga acacgtccct ttcgatgatg tgcatcgcag ccgncattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccctaccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagatcaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gtttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcatag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc gaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940
agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000
tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060
ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120
attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180
tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240
gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300
gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360
caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420
caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480
acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540
atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600
ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660
agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720
atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga 9780
gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840
tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900
gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960
catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020
attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080
tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140
acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200
gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260
gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320
catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380
ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 6
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, WVS

<220>
<221> dấu hiệu hỗn tạp
<222> (763)..(763)
<223> Xaa là alanin hoặc glyxin

<400> 6

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

```
Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
            100                 105                 110

Met Ala Phe His Leu Thr Thr Arg Gly Gly Glu Pro His Met Ile Val
        115                 120                 125

Ser Lys Gln Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly
    130                 135                 140

Val Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Leu Cys Glu
145                 150                 155                 160

Asp Thr Met Thr Tyr Lys Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp
                165                 170                 175

Asp Val Asp Cys Trp Cys Asn Ala Thr Asp Thr Trp Val Thr Tyr Gly
            180                 185                 190

Thr Cys Ser Gln Thr Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
        195                 200                 205

Leu Ala Pro His Val Gly Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp
    210                 215                 220

Met Ser Ser Glu Gly Ala Trp Lys Gln Ile Gln Lys Val Glu Thr Trp
225                 230                 235                 240

Ala Leu Arg His Pro Gly Phe Thr Val Ile Ala Leu Phe Leu Ala His
                245                 250                 255

Ala Ile Gly Thr Ser Ile Thr Gln Lys Gly Ile Ile Phe Ile Leu Leu
            260                 265                 270

Met Leu Val Thr Pro Ser Met Ala Met Arg Cys Val Gly Ile Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
    290                 295                 300
```

```
Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Ile Glu Leu Leu Lys Thr Glu Val Thr Asn Pro Ala Val Leu
                325                 330                 335

Arg Lys Leu Cys Ile Glu Ala Lys Ile Ser Asn Thr Thr Thr Asp Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Ala Thr Leu Val Glu Glu Gln Asp Ala
        355                 360                 365

Asn Phe Val Cys Arg Arg Thr Phe Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Ile Thr Cys Ala Lys Phe Lys
385                 390                 395                 400

Cys Val Thr Lys Leu Glu Gly Lys Ile Val Gln Tyr Glu Asn Leu Lys
                405                 410                 415

Tyr Ser Val Ile Val Thr Val His Thr Gly Asp Gln His Gln Val Gly
            420                 425                 430

Asn Glu Thr Thr Glu His Gly Thr Thr Ala Thr Ile Thr Pro Gln Ala
        435                 440                 445

Pro Thr Ser Glu Ile Gln Leu Thr Asp Tyr Gly Thr Leu Thr Leu Asp
    450                 455                 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Thr
465                 470                 475                 480

Met Lys Glu Arg Ser Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
                485                 490                 495

Pro Leu Pro Trp Thr Ser Gly Ala Ser Thr Ser Gln Glu Thr Trp Asn
            500                 505                 510
```

```
Arg Gln Asp Leu Leu Val Thr Phe Lys Thr Ala His Ala Lys Lys Gln
        515                 520                 525

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

Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Thr Thr Thr Ile Phe Ala
545                 550                 555                 560

Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr Leu Lys Gly
                565                 570                 575

Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys Glu Val
            580                 585                 590

Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr Glu Gly
        595                 600                 605

Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu Lys Gly
    610                 615                 620

Ala Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val Thr Asp
625                 630                 635                 640

Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
                645                 650                 655

Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser Trp Phe
            660                 665                 670

Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly
        675                 680                 685

Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
    690                 695                 700

Ile Gly Gly Val Phe Thr Ser Met Gly Lys Leu Val His Gln Val Phe
705                 710                 715                 720
```

Gly Thr Ala Tyr Gly Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
725 730 735

Ile Gly Ile Gly Ile Leu Leu Thr Trp Leu Gly Leu Asn Ser Arg Asn
740 745 750

Thr Ser Leu Ser Met Met Cys Ile Ala Ala Xaa Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
850 855 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865 870 875 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
885 890 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
915 920 925

```
Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
            980                 985                 990

Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035

Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050

Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065

Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                1075                1080

Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                1090                1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                1105                1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                1120                1125
```

```
His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                1135                1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                1150                1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                1165                1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                1180                1185

Ala Thr  Met Thr Asp Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                1195                1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                1210                1215

Leu Leu  Arg Lys Leu Thr Ser  Lys Glu Leu Met Met  Thr Thr Ile
    1220                1225                1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Leu Pro Glu  Thr Ile Leu
    1235                1240                1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                1255                1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                1270                1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                1285                1290

Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                1300                1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                1315                1320
```

Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu
1325 1330 1335

Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile
1340 1345 1350

Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys
1355 1360 1365

Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu
1370 1375 1380

Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu
1385 1390 1395

Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser
1400 1405 1410

Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser
1415 1420 1425

Met Ser Ile Lys Asn Glu Glu Glu Asp Gln Thr Leu Thr Ile Leu
1430 1435 1440

Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
1445 1450 1455

Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys
1460 1465 1470

Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met
1475 1480 1485

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys
1490 1495 1500

Gly Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu
1505 1510 1515

```
Gly Thr  Phe His Thr Met Trp  His Val Thr Arg Gly  Ala Val Leu
    1520                1525                1530

Met His  Lys Gly Lys Arg Ile  Glu Pro Ser Trp Ala  Asp Val Lys
    1535                1540                1545

Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                1555                1560

Trp Lys  Glu Gly Glu Glu Val  Gln Val Leu Ala Leu  Glu Pro Gly
    1565                1570                1575

Lys Asn  Pro Arg Ala Val Gln  Thr Lys Pro Gly Leu  Phe Lys Thr
    1580                1585                1590

Asn Ala  Gly Thr Ile Gly Ala  Val Ser Leu Asp Phe  Ser Pro Gly
    1595                1600                1605

Thr Ser  Gly Ser Pro Ile Ile  Asp Lys Lys Gly Lys  Val Val Gly
    1610                1615                1620

Leu Tyr  Gly Asn Gly Val Val  Thr Arg Ser Gly Ala  Tyr Val Ser
    1625                1630                1635

Ala Ile  Ala Gln Thr Glu Lys  Ser Ile Glu Asp Asn  Pro Glu Ile
    1640                1645                1650

Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                1660                1665

His Pro  Gly Ala Gly Lys Thr  Lys Arg Tyr Leu Pro  Ala Ile Val
    1670                1675                1680

Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                1690                1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                1705                1710
```

```
Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                 1720                 1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                 1735                 1740

Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                 1750                 1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                 1765                 1770

Tyr Ile  Ser Thr Arg Val Glu  Met Gly Glu Ala Ala  Gly Ile Phe
    1775                 1780                 1785

Met Thr  Ala Thr Pro Pro Gly  Ser Arg Asp Pro Phe  Pro Gln Ser
    1790                 1795                 1800

Asn Ala  Pro Ile Ile Asp Glu  Glu Arg Glu Ile Pro  Glu Arg Ser
    1805                 1810                 1815

Trp Asn  Ser Gly His Glu Trp  Val Thr Asp Phe Lys  Gly Lys Thr
    1820                 1825                 1830

Val Trp  Phe Val Pro Ser Ile  Lys Ala Gly Asn Asp  Ile Ala Ala
    1835                 1840                 1845

Cys Leu  Arg Lys Asn Gly Lys  Lys Val Ile Gln Leu  Ser Arg Lys
    1850                 1855                 1860

Thr Phe  Asp Ser Glu Tyr Val  Lys Thr Arg Thr Asn  Asp Trp Asp
    1865                 1870                 1875

Phe Val  Val Thr Thr Asp Ile  Ser Glu Met Gly Ala  Asn Phe Lys
    1880                 1885                 1890

Ala Glu  Arg Val Ile Asp Pro  Arg Arg Cys Met Lys  Pro Val Ile
    1895                 1900                 1905
```

```
Leu Thr  Asp Gly Glu Glu Arg  Val Ile Leu Ala Gly  Pro Met Pro
    1910                 1915                 1920

Val Thr  His Ser Ser Ala Ala  Gln Arg Arg Gly Arg  Ile Gly Arg
    1925                 1930                 1935

Asn Pro  Lys Asn Glu Asn Asp  Gln Tyr Ile Tyr Met  Gly Glu Pro
    1940                 1945                 1950

Leu Glu  Asn Asp Glu Asp Cys  Ala His Trp Lys Glu  Ala Lys Met
    1955                 1960                 1965

Leu Leu  Asp Asn Ile Asn Thr  Pro Glu Gly Ile Ile  Pro Ser Met
    1970                 1975                 1980

Phe Glu  Pro Glu Arg Glu Lys  Val Asp Ala Ile Asp  Gly Glu Tyr
    1985                 1990                 1995

Arg Leu  Arg Gly Glu Ala Arg  Lys Thr Phe Val Asp  Leu Met Arg
    2000                 2005                 2010

Arg Gly  Asp Leu Pro Val Trp  Leu Ala Tyr Arg Val  Ala Ala Glu
    2015                 2020                 2025

Gly Ile  Asn Tyr Ala Asp Arg  Arg Trp Cys Phe Asp  Gly Val Lys
    2030                 2035                 2040

Asn Asn  Gln Ile Leu Glu Glu  Asn Val Glu Val Glu  Ile Trp Thr
    2045                 2050                 2055

Lys Glu  Gly Glu Arg Lys Lys  Leu Lys Pro Arg Trp  Leu Asp Ala
    2060                 2065                 2070

Arg Ile  Tyr Ser Asp Pro Leu  Ala Leu Lys Glu Phe  Lys Glu Phe
    2075                 2080                 2085

Ala Ala  Gly Arg Lys Ser Leu  Thr Leu Asn Leu Ile  Thr Glu Met
    2090                 2095                 2100
```

Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu
2105 2110 2115

Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala
2120 2125 2130

Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu
2135 2140 2145

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Ile Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg
2525 2530 2535

Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu
2540 2545 2550

Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val
2555 2560 2565

Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly
2570 2575 2580

Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly
2585 2590 2595

Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn
2600 2605 2610

Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro
2615 2620 2625

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp
2810 2815 2820

Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr
2825 2830 2835

Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
2840 2845 2850

Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr
2855 2860 2865

Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg
2870 2875 2880

Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
2885 2890 2895

Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala
2900 2905 2910

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys
2945 2950 2955

Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg
2960 2965 2970

Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp
2975 2980 2985

Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu
2990 2995 3000

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

```
Thr Val  Met Asp Ile Ile Ser  Arg Arg Asp Gln Arg  Gly Ser Gly
    3080                3085                3090

Gln Val  Gly Thr Tyr Gly Leu  Asn Thr Phe Thr Asn  Met Glu Ala
    3095                3100                3105

Gln Leu  Ile Arg Gln Met Glu  Gly Glu Gly Val Phe  Lys Ser Ile
    3110                3115                3120

Gln His  Leu Thr Ile Thr Glu  Glu Ile Ala Val Gln  Asn Trp Leu
    3125                3130                3135

Ala Arg  Val Gly Arg Glu Arg  Leu Ser Arg Met Ala  Ile Ser Gly
    3140                3145                3150

Asp Asp  Cys Val Val Lys Pro  Leu Asp Asp Arg Phe  Ala Ser Ala
    3155                3160                3165

Leu Thr  Ala Leu Asn Asp Met  Gly Lys Ile Arg Lys  Asp Ile Gln
    3170                3175                3180

Gln Trp  Glu Pro Ser Arg Gly  Trp Asn Asp Trp Thr  Gln Val Pro
    3185                3190                3195

Phe Cys  Ser His His Phe His  Glu Leu Ile Met Lys  Asp Gly Arg
    3200                3205                3210

Val Leu  Val Val Pro Cys Arg  Asn Gln Asp Glu Leu  Ile Gly Arg
    3215                3220                3225

Ala Arg  Ile Ser Gln Gly Ala  Gly Trp Ser Leu Arg  Glu Thr Ala
    3230                3235                3240

Cys Leu  Gly Lys Ser Tyr Ala  Gln Met Trp Ser Leu  Met Tyr Phe
    3245                3250                3255

His Arg  Arg Asp Leu Arg Leu  Ala Ala Asn Ala Ile  Cys Ser Ala
    3260                3265                3270
```

Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
3275 3280 3285

His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
3290 3295 3300

Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 7
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, BVS

<400> 7
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgggg gagagccgca tatgatagtt 480

agcaagcagg aaagaggaaa gtcacttttg ttcaagacct ctgcaggtgt caacatgtgc 540

accctcattg cgatggattt gggagagttg tgtgaggaca cgatgaccta caaatgcccc 600

cggatcactg aggcggaacc agatgacgtt gactgttggt gcaatgccac ggacacatgg 660

gtgacctatg gaacgtgctc tcaaactggc gaacaccgac gagacaaacg ttccgtcgca 720

ttggccccac acgtggggct tggcctagaa acaagagccg aaacgtggat gtcctctgaa 780

ggtgcttgga aacagataca aaaagtagag acttgggctc tgagacatcc aggattcacg 840

gtgatagccc tttttctagc acatgccata ggaacatcca tcacccagaa agggatcatt 900

ttcattttgc tgatgctggt aacaccatct atggccatgc gatgcgtggg aataggcaac 960

agagacttcg tggaaggact gtcaggagca acatgggtgg atgtggtact ggagcatgga 1020

agttgcgtca ccaccatggc aaaaaacaaa ccaacactgg acattgaact cttgaagacg 1080

gaggtcacaa accctgcagt tctgcgtaaa ttgtgcattg aagctaaaat atcaaacacc 1140

accaccgatt cgagatgtcc aacacaagga gaagccacac tggtggaaga acaagacgcg 1200

aactttgtgt gccgacgaac gttcgtggac agaggctggg gcaatggctg tgggctattc 1260

ggaaaaggta gtctaataac gtgtgccaag tttaagtgtg tgacaaaact agaaggaaag 1320

atagttcaat atgaaaacct aaaatattca gtgatagtca ccgtccacac tggagatcag 1380

caccaggtgg gaaatgagac tacagaacat ggaacaactg caaccataac acctcaagct 1440

cctacgtcgg aaatacagct gaccgactac ggaacccctta cattagattg ttcacctagg 1500

acagggctag attttaacga gatggtgttg ctgacaatga aagaaagatc atggcttgtc 1560

cacaaacaat ggttcctaga cttaccactg ccttggacct ctggggcttc aacatcccaa 1620

gagacttgga acagacaaga tttactggtc acatttaaga cagctcatgc aaagaagcag 1680

gaagtagtcg tactaggatc acaagaagga gcaatgcaca ctgcgctgac tggagcgaca 1740

gaaatccaaa cgtcaggaac gacaacaatt ttcgcaggac acctaaaatg cagactaaaa 1800

atggacaaac taactttaaa agggatgtca tatgtgatgt gcacaggctc attcaagtta 1860

gagaaagaag tggctgagac ccagcatgga actgttctgg tgcaggttaa atatgaagga 1920

acagacgcac catgcaagat tccctttcg acccaagatg agaaaggagc aacccagaat 1980

gggagattaa taacagccaa ccccatagtc actgacaaag aaaaaccagt caatattgag 2040

gcagaaccac cctttggtga gagctacatc gtggtaggag caggtgaaaa agctttgaaa 2100

ctaagctggt tcaagaaagg aagcagcata gggaaaatgt ttgaagcaac tgcccgagga 2160

gcacgaagga tggccattct gggagacacc gcatgggact tcggttctat aggaggagtg 2220

ttcacgtcta tgggaaaact ggtacaccag gtttttggaa ctgcatatgg agttttgttt 2280

agcggagttt cttggaccat gaaaatagga atagggattc tgctgacatg gctaggatta 2340

aattcaagga acacgtccct ttcgatgatg tgcatcgcag ccgccattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccctaccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagatcaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcatag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc gaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 8
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, BVS

<400> 8

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

```
Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
        35                  40                  45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
    50                  55                  60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65                  70                  75                  80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
                85                  90                  95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
            100                 105                 110

Met Ala Phe His Leu Thr Thr Arg Gly Gly Glu Pro His Met Ile Val
        115                 120                 125

Ser Lys Gln Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly
    130                 135                 140

Val Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Leu Cys Glu
145                 150                 155                 160

Asp Thr Met Thr Tyr Lys Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp
                165                 170                 175

Asp Val Asp Cys Trp Cys Asn Ala Thr Asp Thr Trp Val Thr Tyr Gly
            180                 185                 190

Thr Cys Ser Gln Thr Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
        195                 200                 205

Leu Ala Pro His Val Gly Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp
    210                 215                 220

Met Ser Ser Glu Gly Ala Trp Lys Gln Ile Gln Lys Val Glu Thr Trp
225                 230                 235                 240
```

```
Ala Leu Arg His Pro Gly Phe Thr Val Ile Ala Leu Phe Leu Ala His
                245                 250                 255

Ala Ile Gly Thr Ser Ile Thr Gln Lys Gly Ile Ile Phe Ile Leu Leu
            260                 265                 270

Met Leu Val Thr Pro Ser Met Ala Met Arg Cys Val Gly Ile Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
    290                 295                 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Ile Glu Leu Leu Lys Thr Glu Val Thr Asn Pro Ala Val Leu
                325                 330                 335

Arg Lys Leu Cys Ile Glu Ala Lys Ile Ser Asn Thr Thr Thr Asp Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Ala Thr Leu Val Glu Glu Gln Asp Ala
        355                 360                 365

Asn Phe Val Cys Arg Arg Thr Phe Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Ile Thr Cys Ala Lys Phe Lys
385                 390                 395                 400

Cys Val Thr Lys Leu Glu Gly Lys Ile Val Gln Tyr Glu Asn Leu Lys
                405                 410                 415

Tyr Ser Val Ile Val Thr Val His Thr Gly Asp Gln His Gln Val Gly
            420                 425                 430

Asn Glu Thr Thr Glu His Gly Thr Thr Ala Thr Ile Thr Pro Gln Ala
        435                 440                 445
```

```
Pro Thr Ser Glu Ile Gln Leu Thr Asp Tyr Gly Thr Leu Thr Leu Asp
    450                 455                 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Thr
465                 470                 475                 480

Met Lys Glu Arg Ser Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
                485                 490                 495

Pro Leu Pro Trp Thr Ser Gly Ala Ser Thr Ser Gln Glu Thr Trp Asn
            500                 505                 510

Arg Gln Asp Leu Leu Val Thr Phe Lys Thr Ala His Ala Lys Lys Gln
        515                 520                 525

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

Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Thr Thr Thr Ile Phe Ala
545                 550                 555                 560

Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr Leu Lys Gly
                565                 570                 575

Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys Glu Val
            580                 585                 590

Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr Glu Gly
        595                 600                 605

Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu Lys Gly
    610                 615                 620

Ala Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val Thr Asp
625                 630                 635                 640

Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
                645                 650                 655
```

Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser Trp Phe
660 665 670

Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly
675 680 685

Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
690 695 700

Ile Gly Gly Val Phe Thr Ser Met Gly Lys Leu Val His Gln Val Phe
705 710 715 720

Gly Thr Ala Tyr Gly Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
725 730 735

Ile Gly Ile Gly Ile Leu Leu Thr Trp Leu Gly Leu Asn Ser Arg Asn
740 745 750

Thr Ser Leu Ser Met Met Cys Ile Ala Ala Ala Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
850 855 860

```
Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865                 870                 875                 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
                885                 890                 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
            900                 905                 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
        915                 920                 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
            980                 985                 990

Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035

Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050

Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065
```

```
Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                 1075                 1080

Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                 1090                 1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                 1105                 1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                 1120                 1125

His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                 1135                 1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                 1150                 1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                 1165                 1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                 1180                 1185

Ala Thr  Met Thr Asp Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                 1195                 1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                 1210                 1215

Leu Leu  Arg Lys Leu Thr Ser  Lys Glu Leu Met Met  Thr Thr Ile
    1220                 1225                 1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Leu Pro Glu  Thr Ile Leu
    1235                 1240                 1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                 1255                 1260
```

Val Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala
1265 1270 1275

Ile Leu Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys
1280 1285 1290

Val Ser Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe
1295 1300 1305

Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu
1310 1315 1320

Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu
1325 1330 1335

Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile
1340 1345 1350

Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys
1355 1360 1365

Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu
1370 1375 1380

Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu
1385 1390 1395

Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser
1400 1405 1410

Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser
1415 1420 1425

Met Ser Ile Lys Asn Glu Glu Glu Asp Gln Thr Leu Thr Ile Leu
1430 1435 1440

Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
1445 1450 1455

Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys
1460 1465 1470

Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met
1475 1480 1485

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys
1490 1495 1500

Gly Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu
1505 1510 1515

Gly Thr Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu
1520 1525 1530

Met His Lys Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys
1535 1540 1545

Lys Asp Leu Ile Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu
1550 1555 1560

Trp Lys Glu Gly Glu Glu Val Gln Val Leu Ala Leu Glu Pro Gly
1565 1570 1575

Lys Asn Pro Arg Ala Val Gln Thr Lys Pro Gly Leu Phe Lys Thr
1580 1585 1590

Asn Ala Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser Pro Gly
1595 1600 1605

Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly
1610 1615 1620

Leu Tyr Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser
1625 1630 1635

Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile
1640 1645 1650

```
Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                 1660                 1665

His Pro  Gly Ala Gly Lys Thr  Lys Arg Tyr Leu Pro  Ala Ile Val
    1670                 1675                 1680

Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                 1690                 1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                 1705                 1710

Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                 1720                 1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                 1735                 1740

Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                 1750                 1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                 1765                 1770

Tyr Ile  Ser Thr Arg Val Glu  Met Gly Glu Ala Ala  Gly Ile Phe
    1775                 1780                 1785

Met Thr  Ala Thr Pro Pro Gly  Ser Arg Asp Pro Phe  Pro Gln Ser
    1790                 1795                 1800

Asn Ala  Pro Ile Ile Asp Glu  Glu Arg Glu Ile Pro  Glu Arg Ser
    1805                 1810                 1815

Trp Asn  Ser Gly His Glu Trp  Val Thr Asp Phe Lys  Gly Lys Thr
    1820                 1825                 1830

Val Trp  Phe Val Pro Ser Ile  Lys Ala Gly Asn Asp  Ile Ala Ala
    1835                 1840                 1845
```

```
Cys Leu  Arg Lys Asn Gly Lys  Lys Val Ile Gln Leu  Ser Arg Lys
    1850                 1855                 1860

Thr Phe  Asp Ser Glu Tyr Val  Lys Thr Arg Thr Asn  Asp Trp Asp
    1865                 1870                 1875

Phe Val  Val Thr Thr Asp Ile  Ser Glu Met Gly Ala  Asn Phe Lys
    1880                 1885                 1890

Ala Glu  Arg Val Ile Asp Pro  Arg Arg Cys Met Lys  Pro Val Ile
    1895                 1900                 1905

Leu Thr  Asp Gly Glu Glu Arg  Val Ile Leu Ala Gly  Pro Met Pro
    1910                 1915                 1920

Val Thr  His Ser Ser Ala Ala  Gln Arg Arg Gly Arg  Ile Gly Arg
    1925                 1930                 1935

Asn Pro  Lys Asn Glu Asn Asp  Gln Tyr Ile Tyr Met  Gly Glu Pro
    1940                 1945                 1950

Leu Glu  Asn Asp Glu Asp Cys  Ala His Trp Lys Glu  Ala Lys Met
    1955                 1960                 1965

Leu Leu  Asp Asn Ile Asn Thr  Pro Glu Gly Ile Ile  Pro Ser Met
    1970                 1975                 1980

Phe Glu  Pro Glu Arg Glu Lys  Val Asp Ala Ile Asp  Gly Glu Tyr
    1985                 1990                 1995

Arg Leu  Arg Gly Glu Ala Arg  Lys Thr Phe Val Asp  Leu Met Arg
    2000                 2005                 2010

Arg Gly  Asp Leu Pro Val Trp  Leu Ala Tyr Arg Val  Ala Ala Glu
    2015                 2020                 2025

Gly Ile  Asn Tyr Ala Asp Arg  Arg Trp Cys Phe Asp  Gly Val Lys
    2030                 2035                 2040
```

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met
2090 2095 2100

Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu
2105 2110 2115

Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala
2120 2125 2130

Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu
2135 2140 2145

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Ile Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

```
Cys Glu  Ala Leu Thr Leu Ala  Thr Gly Pro Ile Ser  Thr Leu Trp
    2435                 2440                 2445

Glu Gly  Asn Pro Gly Arg Phe  Trp Asn Thr Thr Ile  Ala Val Ser
    2450                 2455                 2460

Met Ala  Asn Ile Phe Arg Gly  Ser Tyr Leu Ala Gly  Ala Gly Leu
    2465                 2470                 2475

Leu Phe  Ser Ile Met Lys Asn  Thr Thr Asn Thr Arg  Arg Gly Thr
    2480                 2485                 2490

Gly Asn  Ile Gly Glu Thr Leu  Gly Glu Lys Trp Lys  Ser Arg Leu
    2495                 2500                 2505

Asn Ala  Leu Gly Lys Ser Glu  Phe Gln Ile Tyr Lys  Lys Ser Gly
    2510                 2515                 2520

Ile Gln  Glu Val Asp Arg Thr  Leu Ala Lys Glu Gly  Ile Lys Arg
    2525                 2530                 2535

Gly Glu  Thr Asp His His Ala  Val Ser Arg Gly Ser  Ala Lys Leu
    2540                 2545                 2550

Arg Trp  Phe Val Glu Arg Asn  Met Val Thr Pro Glu  Gly Lys Val
    2555                 2560                 2565

Val Asp  Leu Gly Cys Gly Arg  Gly Gly Trp Ser Tyr  Tyr Cys Gly
    2570                 2575                 2580

Gly Leu  Lys Asn Val Arg Glu  Val Lys Gly Leu Thr  Lys Gly Gly
    2585                 2590                 2595

Pro Gly  His Glu Glu Pro Ile  Pro Met Ser Thr Tyr  Gly Trp Asn
    2600                 2605                 2610

Leu Val  Arg Leu Gln Ser Gly  Val Asp Val Phe Phe  Ile Pro Pro
    2615                 2620                 2625
```

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp
2810 2815 2820

Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr
2825 2830 2835

Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
2840 2845 2850

Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr
2855 2860 2865

Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg
2870 2875 2880

Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
2885 2890 2895

Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala
2900 2905 2910

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys
2945 2950 2955

Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg
2960 2965 2970

Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp
2975 2980 2985

Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu
2990 2995 3000

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

```
Gly Ala  Met Tyr Ala Asp Asp  Thr Ala Gly Trp Asp  Thr Arg Ile
    3020                 3025                 3030

Thr Leu  Glu Asp Leu Lys Asn  Glu Glu Met Val Thr  Asn His Met
    3035                 3040                 3045

Glu Gly  Glu His Lys Lys Leu  Ala Glu Ala Ile Phe  Lys Leu Thr
    3050                 3055                 3060

Tyr Gln  Asn Lys Val Val Arg  Val Gln Arg Pro Thr  Pro Arg Gly
    3065                 3070                 3075

Thr Val  Met Asp Ile Ile Ser  Arg Arg Asp Gln Arg  Gly Ser Gly
    3080                 3085                 3090

Gln Val  Gly Thr Tyr Gly Leu  Asn Thr Phe Thr Asn  Met Glu Ala
    3095                 3100                 3105

Gln Leu  Ile Arg Gln Met Glu  Gly Glu Gly Val Phe  Lys Ser Ile
    3110                 3115                 3120

Gln His  Leu Thr Ile Thr Glu  Glu Ile Ala Val Gln  Asn Trp Leu
    3125                 3130                 3135

Ala Arg  Val Gly Arg Glu Arg  Leu Ser Arg Met Ala  Ile Ser Gly
    3140                 3145                 3150

Asp Asp  Cys Val Val Lys Pro  Leu Asp Asp Arg Phe  Ala Ser Ala
    3155                 3160                 3165

Leu Thr  Ala Leu Asn Asp Met  Gly Lys Ile Arg Lys  Asp Ile Gln
    3170                 3175                 3180

Gln Trp  Glu Pro Ser Arg Gly  Trp Asn Asp Trp Thr  Gln Val Pro
    3185                 3190                 3195

Phe Cys  Ser His His Phe His  Glu Leu Ile Met Lys  Asp Gly Arg
    3200                 3205                 3210
```

```
Val Leu  Val Val Pro Cys Arg  Asn Gln Asp Glu Leu  Ile Gly Arg
    3215                 3220                 3225

Ala Arg  Ile Ser Gln Gly Ala  Gly Trp Ser Leu Arg  Glu Thr Ala
    3230                 3235                 3240

Cys Leu  Gly Lys Ser Tyr Ala  Gln Met Trp Ser Leu  Met Tyr Phe
    3245                 3250                 3255

His Arg  Arg Asp Leu Arg Leu  Ala Ala Asn Ala Ile  Cys Ser Ala
    3260                 3265                 3270

Val Pro  Ser His Trp Val Pro  Thr Ser Arg Thr Thr  Trp Ser Ile
    3275                 3280                 3285

His Ala  Lys His Glu Trp Met  Thr Thr Glu Asp Met  Leu Thr Val
    3290                 3295                 3300

Trp Asn  Arg Val Trp Ile Gln  Glu Asn Pro Trp Met  Glu Asp Lys
    3305                 3310                 3315

Thr Pro  Val Glu Ser Trp Glu  Glu Ile Pro Tyr Leu  Gly Lys Arg
    3320                 3325                 3330

Glu Asp  Gln Trp Cys Gly Ser  Leu Ile Gly Leu Thr  Ser Arg Ala
    3335                 3340                 3345

Thr Trp  Ala Lys Asn Ile Gln  Ala Ala Ile Asn Gln  Val Arg Ser
    3350                 3355                 3360

Leu Ile  Gly Asn Glu Glu Tyr  Thr Asp Tyr Met Pro  Ser Met Lys
    3365                 3370                 3375

Arg Phe  Arg Arg Glu Glu Glu  Glu Ala Gly Val Leu  Trp
    3380                 3385                 3390
```

<210> 9
<211> 10723

<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 2, Tiền chủng gốc

<400> 9
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acacgtaacg gagaaccaca catgatcgtc 480

agcagacaag agaaagggaa aagtcttctg tttaaaacag aggttggcgt gaacatgtgt 540

accctcatgg ccatggacct tggtgaattg tgtgaagaca caatcacgta cgagtgtccc 600

cttctcaggc agaatgagcc agaagacata gactgttggt gcaactctac gtccacgtgg 660

gtaacttatg ggacgtgtac caccatggga gaacatagaa gagaaaaaag atcagtggca 720

ctcgttccac atgtgggaat gggactggag acacgaactg aaacatggat gtcatcagaa 780

ggggcctgga aacatgtcca gagaattgaa acttggatct tgagacatcc aggcttcacc 840

atgatggcag caatcctggc atacaccata ggaacgacac atttccaaag agccctgatc 900

ttcatcttac tgacagctgt cactccttca atgacaatgc gttgcatagg aatgtcaaat 960

agagactttg tggaaggggt ttcaggagga agctgggttg acatagtctt agaacatgga 1020

agctgtgtga cgacgatggc aaaaaacaaa ccaacattgg attttgaact gataaaaaca 1080

gaagccaaac agcctgccac cctaaggaag tactgtatag aggcaaagct aaccaacaca 1140

acaacagaat ctcgctgccc aacacaaggg gaacccagcc taaatgaaga gcaggacaaa 1200

aggttcgtct gcaaacactc catggtagac agaggatggg gaaatggatg tggactattt 1260

ggaaagggag gcattgtgac ctgtgctatg ttcagatgca aaaagaacat ggaaggaaaa 1320

gttgtgcaac cagaaaactt ggaatacacc attgtgataa cacctcactc aggggaagag 1380

catgcagtcg gaaatgacac aggaaaacat ggcaaggaaa tcaaaataac accacagagt 1440

tccatcacag aagcagaatt gacaggttat ggcactgtca caatggagtg ctctccaaga 1500

acgggcctcg acttcaatga gatggtgttg ctgcagatgg aaaataaagc ttggctggtg 1560

cacaggcaat ggttcctaga cctgccgtta ccatggttgc ccggagcgga cacacaaggg 1620

tcaaattgga tacagaaaga gacattggtc actttcaaaa atccccatgc gaagaaacag 1680

gatgttgttg ttttaggatc ccaagaaggg gccatgcaca cagcacttac aggggccaca 1740

gaaatccaaa tgtcatcagg aaacttactc ttcacaggac atctcaagtg caggctgaga 1800

atggacaagc tacagctcaa aggaatgtca tactctatgt gcacaggaaa gtttaaagtt 1860

gtgaaggaaa tagcagaaac acaacatgga acaatagtta tcagagtgca atatgaaggg 1920

gacggctctc catgcaagat ccctttttgag ataatggatt tggaaaaaag acatgtctta 1980

ggtcgcctga ttacagtcaa cccaattgtg acagaaaaag atagcccagt caacatagaa 2040

gcagaacctc catttggaga cagctacatc atcataggag tagagccggg acaactgaag 2100

ctcaactggt ttaagaaagg aagttctatc ggccaaatgt ttgagacaac aatgaggggg 2160

gcgaagagaa tggccatttt aggtgacaca gcctgggatt ttggatcctt gggaggagtg 2220

tttacatcta taggaaaggc tctccaccaa gtctttggag caatctatgg agctgccttc 2280

agtggggttt catggactat gaaaatcctc ataggagtca ttatcacatg gataggaatg 2340

aattcacgca gcacctcact gtctgtgaca ctagtattgg tgggaattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctccccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccgtattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 10
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 2, Tiền chủng gốc

<400> 10

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asn Gly Glu Pro His Met Ile Val
115 120 125

Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu Phe Lys Thr Glu Val Gly
130 135 140

Val Asn Met Cys Thr Leu Met Ala Met Asp Leu Gly Glu Leu Cys Glu
145 150 155 160

```
Asp Thr Ile Thr Tyr Glu Cys Pro Leu Leu Arg Gln Asn Glu Pro Glu
                165                 170                 175

Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser Thr Trp Val Thr Tyr Gly
            180                 185                 190

Thr Cys Thr Thr Met Gly Glu His Arg Arg Glu Lys Arg Ser Val Ala
        195                 200                 205

Leu Val Pro His Val Gly Met Gly Leu Glu Thr Arg Thr Glu Thr Trp
    210                 215                 220

Met Ser Ser Glu Gly Ala Trp Lys His Val Gln Arg Ile Glu Thr Trp
225                 230                 235                 240

Ile Leu Arg His Pro Gly Phe Thr Met Met Ala Ala Ile Leu Ala Tyr
                245                 250                 255

Thr Ile Gly Thr Thr His Phe Gln Arg Ala Leu Ile Phe Ile Leu Leu
            260                 265                 270

Thr Ala Val Thr Pro Ser Met Thr Met Arg Cys Ile Gly Met Ser Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ser Trp Val Asp Ile Val
    290                 295                 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala Lys Gln Pro Ala Thr Leu
                325                 330                 335

Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr Asn Thr Thr Thr Glu Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu Asn Glu Glu Gln Asp Lys
        355                 360                 365
```

```
Arg Phe Val Cys Lys His Ser Met Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Gly Ile Val Thr Cys Ala Met Phe Arg
385                 390                 395                 400

Cys Lys Lys Asn Met Glu Gly Lys Val Val Gln Pro Glu Asn Leu Glu
                405                 410                 415

Tyr Thr Ile Val Ile Thr Pro His Ser Gly Glu Glu His Ala Val Gly
            420                 425                 430

Asn Asp Thr Gly Lys His Gly Lys Glu Ile Lys Ile Thr Pro Gln Ser
        435                 440                 445

Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr Gly Thr Val Thr Met Glu
    450                 455                 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Gln
465                 470                 475                 480

Met Glu Asn Lys Ala Trp Leu Val His Arg Gln Trp Phe Leu Asp Leu
                485                 490                 495

Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr Gln Gly Ser Asn Trp Ile
            500                 505                 510

Gln Lys Glu Thr Leu Val Thr Phe Lys Asn Pro His Ala Lys Lys Gln
        515                 520                 525

Asp Val Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu
    530                 535                 540

Thr Gly Ala Thr Glu Ile Gln Met Ser Ser Gly Asn Leu Leu Phe Thr
545                 550                 555                 560

Gly His Leu Lys Cys Arg Leu Arg Met Asp Lys Leu Gln Leu Lys Gly
                565                 570                 575
```

Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val Lys Glu Ile
580 585 590

Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln Tyr Glu Gly
595 600 605

Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu Ile Met Asp Leu Glu Lys
610 615 620

Arg His Val Leu Gly Arg Leu Ile Thr Val Asn Pro Ile Val Thr Glu
625 630 635 640

Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu Asn Trp Phe
660 665 670

Lys Lys Gly Ser Ser Ile Gly Gln Met Phe Glu Thr Thr Met Arg Gly
675 680 685

Ala Lys Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
690 695 700

Leu Gly Gly Val Phe Thr Ser Ile Gly Lys Ala Leu His Gln Val Phe
705 710 715 720

Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly Val Ser Trp Thr Met Lys
725 730 735

Ile Leu Ile Gly Val Ile Ile Thr Trp Ile Gly Met Asn Ser Arg Ser
740 745 750

Thr Ser Leu Ser Val Thr Leu Val Leu Val Gly Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
850 855 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865 870 875 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
885 890 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
915 920 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
930 935 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945 950 955 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
980 985 990

```
Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035

Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050

Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065

Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                1075                1080

Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                1090                1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                1105                1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                1120                1125

His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                1135                1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                1150                1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                1165                1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                1180                1185
```

```
Ala Thr  Met Thr Asp Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                1195                1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                1210                1215

Leu Leu  Arg Lys Leu Thr Ser  Lys Glu Leu Met Met  Thr Thr Ile
    1220                1225                1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Ile Pro Glu  Thr Ile Leu
    1235                1240                1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                1255                1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                1270                1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                1285                1290

Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                1300                1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                1315                1320

Thr Ile  Lys Gly Leu Asn Pro  Thr Ala Ile Phe Leu  Thr Thr Leu
    1325                1330                1335

Ser Arg  Thr Ser Lys Lys Arg  Ser Trp Pro Leu Asn  Glu Ala Ile
    1340                1345                1350

Met Ala  Val Gly Met Val Ser  Ile Leu Ala Ser Ser  Leu Leu Lys
    1355                1360                1365

Asn Asp  Ile Pro Met Thr Gly  Pro Leu Val Ala Gly  Gly Leu Leu
    1370                1375                1380
```

```
Thr Val  Cys Tyr Val Leu Thr  Gly Arg Ser Ala Asp  Leu Glu Leu
    1385                1390                1395

Glu Arg  Ala Ala Asp Val Lys  Trp Glu Asp Gln Ala  Glu Ile Ser
    1400                1405                1410

Gly Ser  Ser Pro Ile Leu Ser  Ile Thr Ile Ser Glu  Asp Gly Ser
    1415                1420                1425

Met Ser  Ile Lys Asn Glu Glu  Glu Glu Gln Thr Leu  Thr Ile Leu
    1430                1435                1440

Ile Arg  Thr Gly Leu Leu Val  Ile Ser Gly Leu Phe  Pro Val Ser
    1445                1450                1455

Ile Pro  Ile Thr Ala Ala Ala  Trp Tyr Leu Trp Glu  Val Lys Lys
    1460                1465                1470

Gln Arg  Ala Gly Val Leu Trp  Asp Val Pro Ser Pro  Pro Pro Met
    1475                1480                1485

Gly Lys  Ala Glu Leu Glu Asp  Gly Ala Tyr Arg Ile  Lys Gln Lys
    1490                1495                1500

Gly Ile  Leu Gly Tyr Ser Gln  Ile Gly Ala Gly Val  Tyr Lys Glu
    1505                1510                1515

Gly Thr  Phe His Thr Met Trp  His Val Thr Arg Gly  Ala Val Leu
    1520                1525                1530

Met His  Lys Gly Lys Arg Ile  Glu Pro Ser Trp Ala  Asp Val Lys
    1535                1540                1545

Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                1555                1560

Trp Lys  Glu Gly Glu Glu Val  Gln Val Leu Ala Leu  Glu Pro Gly
    1565                1570                1575
```

```
Lys Asn  Pro Arg Ala Val Gln  Thr Lys Pro Gly Leu  Phe Lys Thr
    1580                1585                1590

Asn Ala  Gly Thr Ile Gly Ala  Val Ser Leu Asp Phe  Ser Pro Gly
    1595                1600                1605

Thr Ser  Gly Ser Pro Ile Ile  Asp Lys Lys Gly Lys  Val Val Gly
    1610                1615                1620

Leu Tyr  Gly Asn Gly Val Val  Thr Arg Ser Gly Ala  Tyr Val Ser
    1625                1630                1635

Ala Ile  Ala Gln Thr Glu Lys  Ser Ile Glu Asp Asn  Pro Glu Ile
    1640                1645                1650

Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                1660                1665

His Pro  Gly Ala Gly Lys Thr  Lys Arg Tyr Leu Pro  Ala Ile Val
    1670                1675                1680

Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                1690                1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                1705                1710

Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                1720                1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                1735                1740

Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                1750                1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                1765                1770
```

Tyr Ile Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe
1775 1780 1785

Met Thr Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser
1790 1795 1800

Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser
1805 1810 1815

Trp Asn Ser Gly His Glu Trp Val Thr Asp Phe Lys Gly Lys Thr
1820 1825 1830

Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala
1835 1840 1845

Cys Leu Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg Lys
1850 1855 1860

Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp
1865 1870 1875

Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys
1880 1885 1890

Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile
1895 1900 1905

Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro
1910 1915 1920

Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
1925 1930 1935

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
1955 1960 1965

Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met
1970 1975 1980

Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr
1985 1990 1995

Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg
2000 2005 2010

Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu
2015 2020 2025

Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys
2030 2035 2040

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met
2090 2095 2100

Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu
2105 2110 2115

Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala
2120 2125 2130

Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu
2135 2140 2145

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg
2525 2530 2535

Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu
2540 2545 2550

Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val
2555 2560 2565

Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly
2570 2575 2580

Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly
2585 2590 2595

Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn
2600 2605 2610

Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro
2615 2620 2625

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp
2810 2815 2820

Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr
2825 2830 2835

Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
2840 2845 2850

Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr
2855 2860 2865

Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg
2870 2875 2880

Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
2885 2890 2895

Ala Leu Gly Ala Val Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala
2900 2905 2910

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

```
Asn Met  Met Gly Lys Arg Glu  Lys Lys Leu Gly Glu  Phe Gly Lys
    2945                2950                2955

Ala Lys  Gly Ser Arg Ala Ile  Trp Tyr Met Trp Leu  Gly Ala Arg
    2960                2965                2970

Phe Leu  Glu Phe Glu Ala Leu  Gly Phe Leu Asn Glu  Asp His Trp
    2975                2980                2985

Phe Ser  Arg Glu Asn Ser Leu  Ser Gly Val Glu Gly  Glu Gly Leu
    2990                2995                3000

His Lys  Leu Gly Tyr Ile Leu  Arg Asp Val Ser Lys  Lys Glu Gly
    3005                3010                3015

Gly Ala  Met Tyr Ala Asp Asp  Thr Ala Gly Trp Asp  Thr Arg Ile
    3020                3025                3030

Thr Leu  Glu Asp Leu Lys Asn  Glu Glu Met Val Thr  Asn His Met
    3035                3040                3045

Glu Gly  Glu His Lys Lys Leu  Ala Glu Ala Ile Phe  Lys Leu Thr
    3050                3055                3060

Tyr Gln  Asn Lys Val Val Arg  Val Gln Arg Pro Thr  Pro Arg Gly
    3065                3070                3075

Thr Val  Met Asp Ile Ile Ser  Arg Arg Asp Gln Arg  Gly Ser Gly
    3080                3085                3090

Gln Val  Gly Thr Tyr Gly Leu  Asn Thr Phe Thr Asn  Met Glu Ala
    3095                3100                3105

Gln Leu  Ile Arg Gln Met Glu  Gly Glu Gly Val Phe  Lys Ser Ile
    3110                3115                3120

Gln His  Leu Thr Ile Thr Glu  Glu Ile Ala Val Gln  Asn Trp Leu
    3125                3130                3135
```

```
Ala Arg  Val Gly Arg Glu Arg  Leu Ser Arg Met Ala  Ile Ser Gly
    3140                3145                3150

Asp Asp  Cys Val Val Lys Pro  Leu Asp Asp Arg Phe  Ala Ser Ala
    3155                3160                3165

Leu Thr  Ala Leu Asn Asp Met  Gly Lys Ile Arg Lys  Asp Ile Gln
    3170                3175                3180

Gln Trp  Glu Pro Ser Arg Gly  Trp Asn Asp Trp Thr  Gln Val Pro
    3185                3190                3195

Phe Cys  Ser His His Phe His  Glu Leu Ile Met Lys  Asp Gly Arg
    3200                3205                3210

Val Leu  Val Val Pro Cys Arg  Asn Gln Asp Glu Leu  Ile Gly Arg
    3215                3220                3225

Ala Arg  Ile Ser Gln Gly Ala  Gly Trp Ser Leu Arg  Glu Thr Ala
    3230                3235                3240

Cys Leu  Gly Lys Ser Tyr Ala  Gln Met Trp Ser Leu  Met Tyr Phe
    3245                3250                3255

His Arg  Arg Asp Leu Arg Leu  Ala Ala Asn Ala Ile  Cys Ser Ala
    3260                3265                3270

Val Pro  Ser His Trp Val Pro  Thr Ser Arg Thr Thr  Trp Ser Ile
    3275                3280                3285

His Ala  Lys His Glu Trp Met  Thr Thr Glu Asp Met  Leu Thr Val
    3290                3295                3300

Trp Asn  Arg Val Trp Ile Gln  Glu Asn Pro Trp Met  Glu Asp Lys
    3305                3310                3315

Thr Pro  Val Glu Ser Trp Glu  Glu Ile Pro Tyr Leu  Gly Lys Arg
    3320                3325                3330
```

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 11
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 2, MVS

<400> 11
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acacgtaacg gagaaccaca catgatcgtc 480

agcagacaag agaaagggaa aagtcttctg tttaaaacag aggttggcgt gaacatgtgt 540

accctcatgg ccatggacct tggtgaattg tgtgaagaca caatcacgta cgagtgtccc 600

cttctcaggc agaatgagcc agaagacata gactgttggt gcaactctac gtccacgtgg 660

gtaacttatg ggacgtgtac caccatggga gaacatagaa gagaaaaaag atcagtggca 720

ctcgttccac atgtgggaat gggactggag acacgaactg aaacatggat gtcatcagaa 780

ggggcctgga aacatgtcca gagaattgaa acttggatct tgagacatcc aggcttcacc 840

atgatggcag caatcctggc atacaccata ggaacgacac atttccaaag agccctgatc 900

ttcatcttac tgacagctgt cactccttca atgacaatgc gttgcatagg aatgtcaaat 960

agagactttg tggaaggggt ttcaggagga agctgggttg acatagtctt agaacatgga 1020

agctgtgtga cgacgatggc aaaaaacaaa ccaacattgg attttgaact gataaaaaca 1080

gaagccaaac agcctgccac cctaaggaag tactgtatag aggcaaagct aaccaacaca 1140

acaacagaat ctcgctgccc aacacaaggg gaacccagcc taaatgaaga gcaggacaaa 1200

aggttcgtct gcaaacactc catggtagac agaggatggg gaaatggatg tggactattt 1260

ggaaagggag gcattgtgac ctgtgctatg ttcagatgca aaaagaacat ggaaggaaaa 1320

gttgtgcaac cagaaaactt ggaatacacc attgtgataa cacctcactc aggggaagag 1380

catgcagtcg gaaatgacac aggaaaacat ggcaaggaaa tcaaaataac accacagagt 1440

tccatcacag aagcagaatt gacaggttat ggcactgtca caatggagtg ctctccaaga 1500

acgggcctcg acttcaatga gatggtgttg ctgcagatgg aaaataaagc ttggctggtg 1560

cacaggcaat ggttcctaga cctgccgtta ccatggttgc ccggagcgga cacacaaggg 1620

tcaaattgga tacagaaaga gacattggtc actttcaaaa atccccatgc gaagaaacag 1680

gatgttgttg ttttaggatc ccaagaaggg gccatgcaca cagcacttac aggggccaca 1740

gaaatccaaa tgtcatcagg aaacttactc ttcacaggac atctcaagtg caggctgaga 1800

atggacaagc tacagctcaa aggaatgtca tactctatgt gcacaggaaa gtttaaagtt 1860

gtgaaggaaa tagcagaaac acaacatgga acaatagtta tcagagtgca atatgaaggg 1920

gacggctctc catgcaagat ccctttgag ataatggatt tggaaaaaag acatgtctta 1980

ggtcgcctga ttacagtcaa cccaattgtg acagaaaaag atagcccagt caacatagaa 2040

gcagaacctc catttggaga cagctacatc atcataggag tagagccggg acaactgaag 2100

ctcaactggt ttaagaaagg aagttctatc ggccaaatgt ttgagacaac aatgaggggg 2160

gcgaagagaa tggccatttt aggtgacaca gcctgggatt ttggatcctt gggaggagtg 2220

tttacatcta taggaaaggc tctccaccaa gtctttggag caatctatgg agctgccttc 2280

agtggggttt catggactat gaaaatcctc ataggagtca ttatcacatg gataggaatg 2340

aattcacgca gcacctcact gtctgtgaca ctagtattgg tgggaattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gtttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccgtattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 12
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 2, MVS

<400> 12

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asn Gly Glu Pro His Met Ile Val
115 120 125

Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu Phe Lys Thr Glu Val Gly
130 135 140

Val Asn Met Cys Thr Leu Met Ala Met Asp Leu Gly Glu Leu Cys Glu
145 150 155 160

Asp Thr Ile Thr Tyr Glu Cys Pro Leu Leu Arg Gln Asn Glu Pro Glu
165 170 175

Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser Thr Trp Val Thr Tyr Gly
180 185 190

Thr Cys Thr Thr Met Gly Glu His Arg Arg Glu Lys Arg Ser Val Ala
195 200 205

Leu Val Pro His Val Gly Met Gly Leu Glu Thr Arg Thr Glu Thr Trp
210 215 220

Met Ser Ser Glu Gly Ala Trp Lys His Val Gln Arg Ile Glu Thr Trp
225 230 235 240

Ile Leu Arg His Pro Gly Phe Thr Met Met Ala Ala Ile Leu Ala Tyr
245 250 255

Thr Ile Gly Thr Thr His Phe Gln Arg Ala Leu Ile Phe Ile Leu Leu
260 265 270

Thr Ala Val Thr Pro Ser Met Thr Met Arg Cys Ile Gly Met Ser Asn
275 280 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ser Trp Val Asp Ile Val
290 295 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305 310 315 320

Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala Lys Gln Pro Ala Thr Leu
325 330 335

Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr Asn Thr Thr Thr Glu Ser
340 345 350

Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu Asn Glu Glu Gln Asp Lys
355 360 365

Arg Phe Val Cys Lys His Ser Met Val Asp Arg Gly Trp Gly Asn Gly
370 375 380

Cys Gly Leu Phe Gly Lys Gly Gly Ile Val Thr Cys Ala Met Phe Arg
385 390 395 400

Cys Lys Lys Asn Met Glu Gly Lys Val Val Gln Pro Glu Asn Leu Glu
405 410 415

Tyr Thr Ile Val Ile Thr Pro His Ser Gly Glu Glu His Ala Val Gly
420 425 430

Asn Asp Thr Gly Lys His Gly Lys Glu Ile Lys Ile Thr Pro Gln Ser
435 440 445

Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr Gly Thr Val Thr Met Glu
450 455 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Gln
465 470 475 480

Met Glu Asn Lys Ala Trp Leu Val His Arg Gln Trp Phe Leu Asp Leu
485 490 495

Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr Gln Gly Ser Asn Trp Ile
500 505 510

Gln Lys Glu Thr Leu Val Thr Phe Lys Asn Pro His Ala Lys Lys Gln
515 520 525

Asp Val Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu
530 535 540

Thr Gly Ala Thr Glu Ile Gln Met Ser Ser Gly Asn Leu Leu Phe Thr
545 550 555 560

Gly His Leu Lys Cys Arg Leu Arg Met Asp Lys Leu Gln Leu Lys Gly
565 570 575

Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val Lys Glu Ile
580 585 590

Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln Tyr Glu Gly
595 600 605

Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu Ile Met Asp Leu Glu Lys
610 615 620

Arg His Val Leu Gly Arg Leu Ile Thr Val Asn Pro Ile Val Thr Glu
625 630 635 640

Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu Asn Trp Phe
660 665 670

Lys Lys Gly Ser Ser Ile Gly Gln Met Phe Glu Thr Thr Met Arg Gly
675 680 685

Ala Lys Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
690 695 700

Leu Gly Gly Val Phe Thr Ser Ile Gly Lys Ala Leu His Gln Val Phe
705 710 715 720

```
Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly Val Ser Trp Thr Met Lys
                725                 730                 735

Ile Leu Ile Gly Val Ile Ile Thr Trp Ile Gly Met Asn Ser Arg Ser
            740                 745                 750

Thr Ser Leu Ser Val Thr Leu Val Leu Val Gly Ile Val Thr Leu Tyr
        755                 760                 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
    770                 775                 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785                 790                 795                 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
                805                 810                 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
            820                 825                 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
        835                 840                 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
    850                 855                 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865                 870                 875                 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
                885                 890                 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
            900                 905                 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
        915                 920                 925
```

```
Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
            980                 985                 990

Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035

Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050

Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065

Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                1075                1080

Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                1090                1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                1105                1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                1120                1125
```

His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala
1130 1135 1140

Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His
1145 1150 1155

Ala Ile Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly
1160 1165 1170

Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly
1175 1180 1185

Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala
1190 1195 1200

Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu
1205 1210 1215

Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr Ile
1220 1225 1230

Gly Ile Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu
1235 1240 1245

Glu Leu Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met
1250 1255 1260

Val Arg Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala
1265 1270 1275

Ile Leu Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys
1280 1285 1290

Val Ser Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe
1295 1300 1305

Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu
1310 1315 1320

Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu
1325 1330 1335

Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile
1340 1345 1350

Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys
1355 1360 1365

Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu
1370 1375 1380

Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu
1385 1390 1395

Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser
1400 1405 1410

Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser
1415 1420 1425

Met Ser Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu
1430 1435 1440

Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
1445 1450 1455

Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys
1460 1465 1470

Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met
1475 1480 1485

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys
1490 1495 1500

Gly Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu
1505 1510 1515

```
Gly Thr  Phe His Thr Met Trp  His Val Thr Arg Gly  Ala Val Leu
    1520                1525                1530

Met His  Lys Gly Lys Arg Ile  Glu Pro Ser Trp Ala  Asp Val Lys
    1535                1540                1545

Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                1555                1560

Trp Lys  Glu Gly Glu Glu Val  Gln Val Leu Ala Leu  Glu Pro Gly
    1565                1570                1575

Lys Asn  Pro Arg Ala Val Gln  Thr Lys Pro Gly Leu  Phe Lys Thr
    1580                1585                1590

Asn Ala  Gly Thr Ile Gly Ala  Val Ser Leu Asp Phe  Ser Pro Gly
    1595                1600                1605

Thr Ser  Gly Ser Pro Ile Ile  Asp Lys Lys Gly Lys  Val Val Gly
    1610                1615                1620

Leu Tyr  Gly Asn Gly Val Val  Thr Arg Ser Gly Ala  Tyr Val Ser
    1625                1630                1635

Ala Ile  Ala Gln Thr Glu Lys  Ser Ile Glu Asp Asn  Pro Glu Ile
    1640                1645                1650

Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                1660                1665

His Pro  Gly Ala Gly Lys Thr  Lys Arg Tyr Leu Pro  Ala Ile Val
    1670                1675                1680

Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                1690                1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                1705                1710
```

```
Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                 1720                 1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                 1735                 1740

Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                 1750                 1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                 1765                 1770

Tyr Ile  Ser Thr Arg Val Glu  Met Gly Glu Ala Ala  Gly Ile Phe
    1775                 1780                 1785

Met Thr  Ala Thr Pro Pro Gly  Ser Arg Asp Pro Phe  Pro Gln Ser
    1790                 1795                 1800

Asn Ala  Pro Ile Ile Asp Glu  Glu Arg Glu Ile Pro  Glu Arg Ser
    1805                 1810                 1815

Trp Asn  Ser Gly His Glu Trp  Val Thr Asp Phe Lys  Gly Lys Thr
    1820                 1825                 1830

Val Trp  Phe Val Pro Ser Ile  Lys Ala Gly Asn Asp  Ile Ala Ala
    1835                 1840                 1845

Cys Leu  Arg Lys Asn Gly Lys  Lys Val Ile Gln Leu  Ser Arg Lys
    1850                 1855                 1860

Thr Phe  Asp Ser Glu Tyr Val  Lys Thr Arg Thr Asn  Asp Trp Asp
    1865                 1870                 1875

Phe Val  Val Thr Thr Asp Ile  Ser Glu Met Gly Ala  Asn Phe Lys
    1880                 1885                 1890

Ala Glu  Arg Val Ile Asp Pro  Arg Arg Cys Met Lys  Pro Val Ile
    1895                 1900                 1905
```

Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro
1910 1915 1920

Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
1925 1930 1935

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
1955 1960 1965

Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met
1970 1975 1980

Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr
1985 1990 1995

Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg
2000 2005 2010

Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu
2015 2020 2025

Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys
2030 2035 2040

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met
2090 2095 2100

```
Gly Arg  Leu Pro Thr Phe Met  Thr Gln Lys Ala Arg  Asp Ala Leu
    2105                2110                2115

Asp Asn  Leu Ala Val Leu His  Thr Ala Glu Ala Gly  Gly Arg Ala
    2120                2125                2130

Tyr Asn  His Ala Leu Ser Glu  Leu Pro Glu Thr Leu  Glu Thr Leu
    2135                2140                2145

Leu Leu  Leu Thr Leu Leu Ala  Thr Val Thr Gly Gly  Ile Phe Leu
    2150                2155                2160

Phe Leu  Met Ser Ala Arg Gly  Ile Gly Lys Met Thr  Leu Gly Met
    2165                2170                2175

Cys Cys  Ile Ile Thr Ala Ser  Ile Leu Leu Trp Tyr  Ala Gln Ile
    2180                2185                2190

Gln Pro  His Trp Ile Ala Ala  Ser Ile Ile Leu Glu  Phe Phe Leu
    2195                2200                2205

Ile Val  Leu Leu Ile Pro Glu  Pro Glu Lys Gln Arg  Thr Pro Gln
    2210                2215                2220

Asp Asn  Gln Leu Thr Tyr Val  Val Ile Ala Ile Leu  Thr Val Val
    2225                2230                2235

Ala Ala  Thr Met Ala Asn Glu  Met Gly Phe Leu Glu  Lys Thr Lys
    2240                2245                2250

Lys Asp  Leu Gly Leu Gly Ser  Ile Ala Thr Gln Gln  Pro Glu Ser
    2255                2260                2265

Asn Ile  Leu Asp Ile Asp Leu  Arg Pro Ala Ser Ala  Trp Thr Leu
    2270                2275                2280

Tyr Ala  Val Ala Thr Thr Phe  Val Thr Pro Met Leu  Arg His Ser
    2285                2290                2295
```

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg
2525 2530 2535

Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu
2540 2545 2550

Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val
2555 2560 2565

Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly
2570 2575 2580

Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly
2585 2590 2595

Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn
2600 2605 2610

Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro
2615 2620 2625

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp
2810 2815 2820

Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr
2825 2830 2835

Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
2840 2845 2850

Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr
2855 2860 2865

Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg
2870 2875 2880

```
Met Cys  Thr Arg Glu Glu Phe  Thr Arg Lys Val Arg  Ser Asn Ala
    2885                 2890                 2895

Ala Leu  Gly Ala Val Phe Thr  Asp Glu Asn Lys Trp  Lys Ser Ala
    2900                 2905                 2910

Arg Glu  Ala Val Glu Asp Ser  Arg Phe Trp Glu Leu  Val Asp Lys
    2915                 2920                 2925

Glu Arg  Asn Leu His Leu Glu  Gly Lys Cys Glu Thr  Cys Val Tyr
    2930                 2935                 2940

Asn Met  Met Gly Lys Arg Glu  Lys Lys Leu Gly Glu  Phe Gly Lys
    2945                 2950                 2955

Ala Lys  Gly Ser Arg Ala Ile  Trp Tyr Met Trp Leu  Gly Ala Arg
    2960                 2965                 2970

Phe Leu  Glu Phe Glu Ala Leu  Gly Phe Leu Asn Glu  Asp His Trp
    2975                 2980                 2985

Phe Ser  Arg Glu Asn Ser Leu  Ser Gly Val Glu Gly  Glu Gly Leu
    2990                 2995                 3000

His Lys  Leu Gly Tyr Ile Leu  Arg Asp Val Ser Lys  Lys Glu Gly
    3005                 3010                 3015

Gly Ala  Met Tyr Ala Asp Asp  Thr Ala Gly Trp Asp  Thr Arg Ile
    3020                 3025                 3030

Thr Leu  Glu Asp Leu Lys Asn  Glu Glu Met Val Thr  Asn His Met
    3035                 3040                 3045

Glu Gly  Glu His Lys Lys Leu  Ala Glu Ala Ile Phe  Lys Leu Thr
    3050                 3055                 3060

Tyr Gln  Asn Lys Val Val Arg  Val Gln Arg Pro Thr  Pro Arg Gly
    3065                 3070                 3075
```

Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
3080 3085 3090

Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala
3095 3100 3105

Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile
3110 3115 3120

Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
3125 3130 3135

Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly
3140 3145 3150

Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
3155 3160 3165

Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln
3170 3175 3180

Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro
3185 3190 3195

Phe Cys Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg
3200 3205 3210

Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg
3215 3220 3225

Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala
3230 3235 3240

Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe
3245 3250 3255

His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
3260 3265 3270

```
Val Pro  Ser His Trp Val Pro  Thr Ser Arg Thr Thr  Trp Ser Ile
    3275                 3280                 3285

His Ala  Lys His Glu Trp Met  Thr Thr Glu Asp Met  Leu Thr Val
    3290                 3295                 3300

Trp Asn  Arg Val Trp Ile Gln  Glu Asn Pro Trp Met  Glu Asp Lys
    3305                 3310                 3315

Thr Pro  Val Glu Ser Trp Glu  Glu Ile Pro Tyr Leu  Gly Lys Arg
    3320                 3325                 3330

Glu Asp  Gln Trp Cys Gly Ser  Leu Ile Gly Leu Thr  Ser Arg Ala
    3335                 3340                 3345

Thr Trp  Ala Lys Asn Ile Gln  Ala Ala Ile Asn Gln  Val Arg Ser
    3350                 3355                 3360

Leu Ile  Gly Asn Glu Glu Tyr  Thr Asp Tyr Met Pro  Ser Met Lys
    3365                 3370                 3375

Arg Phe  Arg Arg Glu Glu Glu  Glu Ala Gly Val Leu  Trp
    3380                 3385                 3390
```

<210> 13
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 2, WVS

<220>
<221> dấu hiệu hỗn tạp
<222> (6481)..(6481)
<223> n là xytosin hoặc guanin

<220>
<221> dấu hiệu hỗn tạp
<222> (7156)..(7156)
<223> n là xytosin hoặc thymin

<400> 13

agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acacgtaacg gagaaccaca catgatcgtc 480

agcagacaag agaaagggaa aagtcttctg tttaaaacag aggttggcgt gaacatgtgt 540

accctcatgg ccatggacct tggtgaattg tgtgaagaca caatcacgta cgagtgtccc 600

cttctcaggc agaatgagcc agaagacata gactgttggt gcaactctac gtccacgtgg 660

gtaacttatg ggacgtgtac caccatggga gaacatagaa gagaaaaaag atcagtggca 720

ctcgttccac atgtgggaat gggactggag acacgaactg aaacatggat gtcatcagaa 780

ggggcctgga aacatgtcca gagaattgaa acttggatct tgagacatcc aggcttcacc 840

atgatggcag caatcctggc atacaccata ggaacgacac atttccaaag agccctgatc 900

ttcatcttac tgacagctgt cactccttca atgacaatgc gttgcatagg aatgtcaaat 960

agagactttg tggaaggggt ttcaggagga agctgggttg acatagtctt agaacatgga 1020

agctgtgtga cgacgatggc aaaaaacaaa ccaacattgg attttgaact gataaaaaca 1080

gaagccaaac agcctgccac cctaaggaag tactgtatag aggcaaagct aaccaacaca 1140

acaacagaat ctcgctgccc aacacaaggg gaacccagcc taaatgaaga gcaggacaaa 1200

aggttcgtct gcaaacactc catggtagac agaggatggg gaaatggatg tggactattt 1260

ggaaagggag gcattgtgac ctgtgctatg ttcagatgca aaaagaacat ggaaggaaaa 1320

gttgtgcaac cagaaaactt ggaatacacc attgtgataa cacctcactc aggggaagag 1380

catgcagtcg gaaatgacac aggaaaacat ggcaaggaaa tcaaaataac accacagagt 1440

tccatcacag aagcagaatt gacaggttat ggcactgtca caatggagtg ctctccaaga 1500

acgggcctcg acttcaatga gatggtgttg ctgcagatgg aaaataaagc ttggctggtg 1560

cacaggcaat ggttcctaga cctgccgtta ccatggttgc ccggagcgga cacacaaggg 1620

tcaaattgga tacagaaaga gacattggtc actttcaaaa atccccatgc gaagaaacag 1680

gatgttgttg ttttaggatc ccaagaaggg gccatgcaca cagcacttac aggggccaca 1740

gaaatccaaa tgtcatcagg aaacttactc ttcacaggac atctcaagtg caggctgaga 1800

atggacaagc tacagctcaa aggaatgtca tactctatgt gcacaggaaa gtttaaagtt 1860

gtgaaggaaa tagcagaaac acaacatgga acaatagtta tcagagtgca atatgaaggg 1920

gacggctctc catgcaagat ccctttgag ataatggatt tggaaaaaag acatgtctta 1980

ggtcgcctga ttacagtcaa cccaattgtg acagaaaaag atagcccagt caacatagaa 2040

gcagaacctc catttggaga cagctacatc atcataggag tagagccggg acaactgaag 2100

ctcaactggt ttaagaaagg aagttctatc ggccaaatgt ttgagacaac aatgaggggg 2160

gcgaagagaa tggccatttt aggtgacaca gcctgggatt ttggatcctt gggaggagtg 2220

tttacatcta taggaaaggc tctccaccaa gtctttggag caatctatgg agctgccttc 2280

agtggggttt catggactat gaaaatcctc ataggagtca ttatcacatg gataggaatg 2340

aattcacgca gcacctcact gtctgtgaca ctagtattgg tgggaattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

ncaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gtttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctntttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccgtattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 14
<211> 3391
<212> PRT

<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 2, WVS

<220>
<221> dấu hiệu hỗn tạp
<222> (2129)..(2129)
<223> Xaa là alanin hoặc prolin

<220>
<221> dấu hiệu hỗn tạp
<222> (2354)..(2354)
<223> Xaa là leuxin hoặc phenylalanin

<400> 14

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asn Gly Glu Pro His Met Ile Val
115 120 125

Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu Phe Lys Thr Glu Val Gly
130 135 140

Val Asn Met Cys Thr Leu Met Ala Met Asp Leu Gly Glu Leu Cys Glu
145 150 155 160

Asp Thr Ile Thr Tyr Glu Cys Pro Leu Leu Arg Gln Asn Glu Pro Glu
165 170 175

Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser Thr Trp Val Thr Tyr Gly
180 185 190

Thr Cys Thr Thr Met Gly Glu His Arg Arg Glu Lys Arg Ser Val Ala
195 200 205

Leu Val Pro His Val Gly Met Gly Leu Glu Thr Arg Thr Glu Thr Trp
210 215 220

Met Ser Ser Glu Gly Ala Trp Lys His Val Gln Arg Ile Glu Thr Trp
225 230 235 240

Ile Leu Arg His Pro Gly Phe Thr Met Met Ala Ala Ile Leu Ala Tyr
245 250 255

Thr Ile Gly Thr Thr His Phe Gln Arg Ala Leu Ile Phe Ile Leu Leu
260 265 270

Thr Ala Val Thr Pro Ser Met Thr Met Arg Cys Ile Gly Met Ser Asn
275 280 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ser Trp Val Asp Ile Val
290 295 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305 310 315 320

Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala Lys Gln Pro Ala Thr Leu
325 330 335

Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr Asn Thr Thr Thr Glu Ser
340 345 350

```
Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu Asn Glu Glu Gln Asp Lys
        355                 360                 365

Arg Phe Val Cys Lys His Ser Met Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Gly Ile Val Thr Cys Ala Met Phe Arg
385                 390                 395                 400

Cys Lys Lys Asn Met Glu Gly Lys Val Val Gln Pro Glu Asn Leu Glu
                405                 410                 415

Tyr Thr Ile Val Ile Thr Pro His Ser Gly Glu Glu His Ala Val Gly
            420                 425                 430

Asn Asp Thr Gly Lys His Gly Lys Glu Ile Lys Ile Thr Pro Gln Ser
        435                 440                 445

Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr Gly Thr Val Thr Met Glu
    450                 455                 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Gln
465                 470                 475                 480

Met Glu Asn Lys Ala Trp Leu Val His Arg Gln Trp Phe Leu Asp Leu
                485                 490                 495

Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr Gln Gly Ser Asn Trp Ile
            500                 505                 510

Gln Lys Glu Thr Leu Val Thr Phe Lys Asn Pro His Ala Lys Lys Gln
        515                 520                 525

Asp Val Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu
    530                 535                 540

Thr Gly Ala Thr Glu Ile Gln Met Ser Ser Gly Asn Leu Leu Phe Thr
545                 550                 555                 560
```

Gly His Leu Lys Cys Arg Leu Arg Met Asp Lys Leu Gln Leu Lys Gly
565 570 575

Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val Lys Glu Ile
580 585 590

Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln Tyr Glu Gly
595 600 605

Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu Ile Met Asp Leu Glu Lys
610 615 620

Arg His Val Leu Gly Arg Leu Ile Thr Val Asn Pro Ile Val Thr Glu
625 630 635 640

Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu Asn Trp Phe
660 665 670

Lys Lys Gly Ser Ser Ile Gly Gln Met Phe Glu Thr Thr Met Arg Gly
675 680 685

Ala Lys Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
690 695 700

Leu Gly Gly Val Phe Thr Ser Ile Gly Lys Ala Leu His Gln Val Phe
705 710 715 720

Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly Val Ser Trp Thr Met Lys
725 730 735

Ile Leu Ile Gly Val Ile Ile Thr Trp Ile Gly Met Asn Ser Arg Ser
740 745 750

Thr Ser Leu Ser Val Thr Leu Val Leu Val Gly Ile Val Thr Leu Tyr
755 760 765

```
Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
    770                 775                 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785                 790                 795                 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
                805                 810                 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
            820                 825                 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
        835                 840                 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
    850                 855                 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865                 870                 875                 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
                885                 890                 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
            900                 905                 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
        915                 920                 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975
```

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
1025 1030 1035

Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp
1040 1045 1050

Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn
1055 1060 1065

Arg Gly Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile
1070 1075 1080

Thr Glu Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr
1085 1090 1095

Arg Gly Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu
1100 1105 1110

Lys Glu Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly
1115 1120 1125

His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala
1130 1135 1140

Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His
1145 1150 1155

Ala Ile Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly
1160 1165 1170

```
Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                 1180                 1185

Ala Thr  Met Thr Asp Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                 1195                 1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                 1210                 1215

Leu Leu  Arg Lys Leu Thr Ser  Lys Glu Leu Met Met  Thr Thr Ile
    1220                 1225                 1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Ile Pro Glu  Thr Ile Leu
    1235                 1240                 1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                 1255                 1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                 1270                 1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                 1285                 1290

Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                 1300                 1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                 1315                 1320

Thr Ile  Lys Gly Leu Asn Pro  Thr Ala Ile Phe Leu  Thr Thr Leu
    1325                 1330                 1335

Ser Arg  Thr Ser Lys Lys Arg  Ser Trp Pro Leu Asn  Glu Ala Ile
    1340                 1345                 1350

Met Ala  Val Gly Met Val Ser  Ile Leu Ala Ser Ser  Leu Leu Lys
    1355                 1360                 1365
```

```
Asn Asp  Ile Pro Met Thr Gly  Pro Leu Val Ala Gly  Gly Leu Leu
    1370                 1375                 1380

Thr Val  Cys Tyr Val Leu Thr  Gly Arg Ser Ala Asp  Leu Glu Leu
    1385                 1390                 1395

Glu Arg  Ala Ala Asp Val Lys  Trp Glu Asp Gln Ala  Glu Ile Ser
    1400                 1405                 1410

Gly Ser  Ser Pro Ile Leu Ser  Ile Thr Ile Ser Glu  Asp Gly Ser
    1415                 1420                 1425

Met Ser  Ile Lys Asn Glu Glu  Glu Glu Gln Thr Leu  Thr Ile Leu
    1430                 1435                 1440

Ile Arg  Thr Gly Leu Leu Val  Ile Ser Gly Leu Phe  Pro Val Ser
    1445                 1450                 1455

Ile Pro  Ile Thr Ala Ala Ala  Trp Tyr Leu Trp Glu  Val Lys Lys
    1460                 1465                 1470

Gln Arg  Ala Gly Val Leu Trp  Asp Val Pro Ser Pro  Pro Pro Met
    1475                 1480                 1485

Gly Lys  Ala Glu Leu Glu Asp  Gly Ala Tyr Arg Ile  Lys Gln Lys
    1490                 1495                 1500

Gly Ile  Leu Gly Tyr Ser Gln  Ile Gly Ala Gly Val  Tyr Lys Glu
    1505                 1510                 1515

Gly Thr  Phe His Thr Met Trp  His Val Thr Arg Gly  Ala Val Leu
    1520                 1525                 1530

Met His  Lys Gly Lys Arg Ile  Glu Pro Ser Trp Ala  Asp Val Lys
    1535                 1540                 1545

Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                 1555                 1560
```

Trp Lys Glu Gly Glu Glu Val Gln Val Leu Ala Leu Glu Pro Gly
1565 1570 1575

Lys Asn Pro Arg Ala Val Gln Thr Lys Pro Gly Leu Phe Lys Thr
1580 1585 1590

Asn Ala Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser Pro Gly
1595 1600 1605

Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly
1610 1615 1620

Leu Tyr Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser
1625 1630 1635

Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile
1640 1645 1650

Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu
1655 1660 1665

His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val
1670 1675 1680

Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro
1685 1690 1695

Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu
1700 1705 1710

Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly
1715 1720 1725

Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg
1730 1735 1740

Leu Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met
1745 1750 1755

Asp Glu Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly
1760 1765 1770

Tyr Ile Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe
1775 1780 1785

Met Thr Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser
1790 1795 1800

Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser
1805 1810 1815

Trp Asn Ser Gly His Glu Trp Val Thr Asp Phe Lys Gly Lys Thr
1820 1825 1830

Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala
1835 1840 1845

Cys Leu Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg Lys
1850 1855 1860

Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp
1865 1870 1875

Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys
1880 1885 1890

Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile
1895 1900 1905

Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro
1910 1915 1920

Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
1925 1930 1935

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

```
Leu Glu  Asn Asp Glu Asp Cys  Ala His Trp Lys Glu  Ala Lys Met
    1955                 1960                 1965

Leu Leu  Asp Asn Ile Asn Thr  Pro Glu Gly Ile Ile  Pro Ser Met
    1970                 1975                 1980

Phe Glu  Pro Glu Arg Glu Lys  Val Asp Ala Ile Asp  Gly Glu Tyr
    1985                 1990                 1995

Arg Leu  Arg Gly Glu Ala Arg  Lys Thr Phe Val Asp  Leu Met Arg
    2000                 2005                 2010

Arg Gly  Asp Leu Pro Val Trp  Leu Ala Tyr Arg Val  Ala Ala Glu
    2015                 2020                 2025

Gly Ile  Asn Tyr Ala Asp Arg  Arg Trp Cys Phe Asp  Gly Val Lys
    2030                 2035                 2040

Asn Asn  Gln Ile Leu Glu Glu  Asn Val Glu Val Glu  Ile Trp Thr
    2045                 2050                 2055

Lys Glu  Gly Glu Arg Lys Lys  Leu Lys Pro Arg Trp  Leu Asp Ala
    2060                 2065                 2070

Arg Ile  Tyr Ser Asp Pro Leu  Ala Leu Lys Glu Phe  Lys Glu Phe
    2075                 2080                 2085

Ala Ala  Gly Arg Lys Ser Leu  Thr Leu Asn Leu Ile  Thr Glu Met
    2090                 2095                 2100

Gly Arg  Leu Pro Thr Phe Met  Thr Gln Lys Ala Arg  Asp Ala Leu
    2105                 2110                 2115

Asp Asn  Leu Ala Val Leu His  Thr Ala Glu Xaa Gly  Gly Arg Ala
    2120                 2125                 2130

Tyr Asn  His Ala Leu Ser Glu  Leu Pro Glu Thr Leu  Glu Thr Leu
    2135                 2140                 2145
```

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Xaa Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg
2525 2530 2535

```
Gly Glu  Thr Asp His His Ala  Val Ser Arg Gly Ser  Ala Lys Leu
    2540                2545                2550

Arg Trp  Phe Val Glu Arg Asn  Met Val Thr Pro Glu  Gly Lys Val
    2555                2560                2565

Val Asp  Leu Gly Cys Gly Arg  Gly Gly Trp Ser Tyr  Tyr Cys Gly
    2570                2575                2580

Gly Leu  Lys Asn Val Arg Glu  Val Lys Gly Leu Thr  Lys Gly Gly
    2585                2590                2595

Pro Gly  His Glu Glu Pro Ile  Pro Met Ser Thr Tyr  Gly Trp Asn
    2600                2605                2610

Leu Val  Arg Leu Gln Ser Gly  Val Asp Val Phe Phe  Ile Pro Pro
    2615                2620                2625

Glu Lys  Cys Asp Thr Leu Leu  Cys Asp Ile Gly Glu  Ser Ser Pro
    2630                2635                2640

Asn Pro  Thr Val Glu Ala Gly  Arg Thr Leu Arg Val  Leu Asn Leu
    2645                2650                2655

Val Glu  Asn Trp Leu Asn Asn  Asn Thr Gln Phe Cys  Ile Lys Val
    2660                2665                2670

Leu Asn  Pro Tyr Met Pro Ser  Val Ile Glu Lys Met  Glu Ala Leu
    2675                2680                2685

Gln Arg  Lys Tyr Gly Gly Ala  Leu Val Arg Asn Pro  Leu Ser Arg
    2690                2695                2700

Asn Ser  Thr His Glu Met Tyr  Trp Val Ser Asn Ala  Ser Gly Asn
    2705                2710                2715

Ile Val  Ser Ser Val Asn Met  Ile Ser Arg Met Leu  Ile Asn Arg
    2720                2725                2730
```

```
Phe Thr  Met Arg Tyr Lys Lys  Ala Thr Tyr Glu Pro  Asp Val Asp
    2735                2740                2745

Leu Gly  Ser Gly Thr Arg Asn  Ile Gly Ile Glu Ser  Glu Ile Pro
    2750                2755                2760

Asn Leu  Asp Ile Ile Gly Lys  Arg Ile Glu Lys Ile  Lys Gln Glu
    2765                2770                2775

His Glu  Thr Ser Trp His Tyr  Asp Gln Asp His Pro  Tyr Lys Thr
    2780                2785                2790

Trp Ala  Tyr His Gly Ser Tyr  Glu Thr Lys Gln Thr  Gly Ser Ala
    2795                2800                2805

Ser Ser  Met Val Asn Gly Val  Val Arg Leu Leu Thr  Lys Pro Trp
    2810                2815                2820

Asp Val  Val Pro Met Val Thr  Gln Met Ala Met Thr  Asp Thr Thr
    2825                2830                2835

Pro Phe  Gly Gln Gln Arg Val  Phe Lys Glu Lys Val  Asp Thr Arg
    2840                2845                2850

Thr Gln  Glu Pro Lys Glu Gly  Thr Lys Lys Leu Met  Lys Ile Thr
    2855                2860                2865

Ala Glu  Trp Leu Trp Lys Glu  Leu Gly Lys Lys Lys  Thr Pro Arg
    2870                2875                2880

Met Cys  Thr Arg Glu Glu Phe  Thr Arg Lys Val Arg  Ser Asn Ala
    2885                2890                2895

Ala Leu  Gly Ala Val Phe Thr  Asp Glu Asn Lys Trp  Lys Ser Ala
    2900                2905                2910

Arg Glu  Ala Val Glu Asp Ser  Arg Phe Trp Glu Leu  Val Asp Lys
    2915                2920                2925
```

```
Glu Arg  Asn Leu His Leu Glu  Gly Lys Cys Glu Thr  Cys Val Tyr
    2930                 2935                 2940

Asn Met  Met Gly Lys Arg Glu  Lys Lys Leu Gly Glu  Phe Gly Lys
    2945                 2950                 2955

Ala Lys  Gly Ser Arg Ala Ile  Trp Tyr Met Trp Leu  Gly Ala Arg
    2960                 2965                 2970

Phe Leu  Glu Phe Glu Ala Leu  Gly Phe Leu Asn Glu  Asp His Trp
    2975                 2980                 2985

Phe Ser  Arg Glu Asn Ser Leu  Ser Gly Val Glu Gly  Glu Gly Leu
    2990                 2995                 3000

His Lys  Leu Gly Tyr Ile Leu  Arg Asp Val Ser Lys  Lys Glu Gly
    3005                 3010                 3015

Gly Ala  Met Tyr Ala Asp Asp  Thr Ala Gly Trp Asp  Thr Arg Ile
    3020                 3025                 3030

Thr Leu  Glu Asp Leu Lys Asn  Glu Glu Met Val Thr  Asn His Met
    3035                 3040                 3045

Glu Gly  Glu His Lys Lys Leu  Ala Glu Ala Ile Phe  Lys Leu Thr
    3050                 3055                 3060

Tyr Gln  Asn Lys Val Val Arg  Val Gln Arg Pro Thr  Pro Arg Gly
    3065                 3070                 3075

Thr Val  Met Asp Ile Ile Ser  Arg Arg Asp Gln Arg  Gly Ser Gly
    3080                 3085                 3090

Gln Val  Gly Thr Tyr Gly Leu  Asn Thr Phe Thr Asn  Met Glu Ala
    3095                 3100                 3105

Gln Leu  Ile Arg Gln Met Glu  Gly Glu Gly Val Phe  Lys Ser Ile
    3110                 3115                 3120
```

```
Gln His  Leu Thr Ile Thr Glu  Glu Ile Ala Val Gln  Asn Trp Leu
    3125                3130                3135

Ala Arg  Val Gly Arg Glu Arg  Leu Ser Arg Met Ala  Ile Ser Gly
    3140                3145                3150

Asp Asp  Cys Val Val Lys Pro  Leu Asp Asp Arg Phe  Ala Ser Ala
    3155                3160                3165

Leu Thr  Ala Leu Asn Asp Met  Gly Lys Ile Arg Lys  Asp Ile Gln
    3170                3175                3180

Gln Trp  Glu Pro Ser Arg Gly  Trp Asn Asp Trp Thr  Gln Val Pro
    3185                3190                3195

Phe Cys  Ser His His Phe His  Glu Leu Ile Met Lys  Asp Gly Arg
    3200                3205                3210

Val Leu  Val Val Pro Cys Arg  Asn Gln Asp Glu Leu  Ile Gly Arg
    3215                3220                3225

Ala Arg  Ile Ser Gln Gly Ala  Gly Trp Ser Leu Arg  Glu Thr Ala
    3230                3235                3240

Cys Leu  Gly Lys Ser Tyr Ala  Gln Met Trp Ser Leu  Met Tyr Phe
    3245                3250                3255

His Arg  Arg Asp Leu Arg Leu  Ala Ala Asn Ala Ile  Cys Ser Ala
    3260                3265                3270

Val Pro  Ser His Trp Val Pro  Thr Ser Arg Thr Thr  Trp Ser Ile
    3275                3280                3285

His Ala  Lys His Glu Trp Met  Thr Thr Glu Asp Met  Leu Thr Val
    3290                3295                3300

Trp Asn  Arg Val Trp Ile Gln  Glu Asn Pro Trp Met  Glu Asp Lys
    3305                3310                3315
```

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 15
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 2, BVS

<400> 15
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acacgtaacg gagaaccaca catgatcgtc 480

agcagacaag agaaagggaa aagtcttctg tttaaaacag aggttggcgt gaacatgtgt 540

accctcatgg ccatggacct tggtgaattg tgtgaagaca caatcacgta cgagtgtccc 600

cttctcaggc agaatgagcc agaagacata gactgttggt gcaactctac gtccacgtgg 660

gtaacttatg ggacgtgtac caccatggga gaacatagaa gagaaaaaag atcagtggca 720

ctcgttccac atgtgggaat gggactggag acacgaactg aaacatggat gtcatcagaa 780

ggggcctgga aacatgtcca gagaattgaa acttggatct tgagacatcc aggcttcacc 840

atgatggcag caatcctggc atacaccata ggaacgacac atttccaaag agccctgatc 900

ttcatcttac tgacagctgt cactccttca atgacaatgc gttgcatagg aatgtcaaat 960

agagactttg tggaaggggt ttcaggagga agctgggttg acatagtctt agaacatgga 1020

agctgtgtga cgacgatggc aaaaaacaaa ccaacattgg attttgaact gataaaaaca 1080

gaagccaaac agcctgccac cctaaggaag tactgtatag aggcaaagct aaccaacaca 1140

acaacagaat ctcgctgccc aacacaaggg gaacccagcc taaatgaaga gcaggacaaa 1200

aggttcgtct gcaaacactc catggtagac agaggatggg gaaatggatg tggactattt 1260

ggaaagggag gcattgtgac ctgtgctatg ttcagatgca aaaagaacat ggaaggaaaa 1320

gttgtgcaac cagaaaactt ggaatacacc attgtgataa cacctcactc aggggaagag 1380

catgcagtcg gaaatgacac aggaaaacat ggcaaggaaa tcaaaataac accacagagt 1440

tccatcacag aagcagaatt gacaggttat ggcactgtca caatggagtg ctctccaaga 1500

acgggcctcg acttcaatga gatggtgttg ctgcagatgg aaaataaagc ttggctggtg 1560

cacaggcaat ggttcctaga cctgccgtta ccatggttgc ccggagcgga cacacaaggg 1620

tcaaattgga tacagaaaga gacattggtc actttcaaaa atccccatgc gaagaaacag 1680

gatgttgttg ttttaggatc ccaagaaggg gccatgcaca cagcacttac aggggccaca 1740

gaaatccaaa tgtcatcagg aaacttactc ttcacaggac atctcaagtg caggctgaga 1800

atggacaagc tacagctcaa aggaatgtca tactctatgt gcacaggaaa gtttaaagtt 1860

gtgaaggaaa tagcagaaac acaacatgga acaatagtta tcagagtgca atatgaaggg 1920

gacggctctc catgcaagat ccctttgag ataatggatt tggaaaaaag acatgtctta 1980

ggtcgcctga ttacagtcaa cccaattgtg acagaaaaag atagcccagt caacatagaa 2040

gcagaacctc catttggaga cagctacatc atcataggag tagagccggg acaactgaag 2100

ctcaactggt ttaagaaagg aagttctatc ggccaaatgt ttgagacaac aatgaggggg 2160

gcgaagagaa tggccatttt aggtgacaca gcctgggatt ttggatcctt gggaggagtg 2220

tttacatcta taggaaaggc tctccaccaa gtctttggag caatctatgg agctgccttc 2280

agtggggttt catggactat gaaaatcctc ataggagtca ttatcacatg gataggaatg 2340

aattcacgca gcacctcact gtctgtgaca ctagtattgg tgggaattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctccccgggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

ccaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gtttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctttttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccgtattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 16
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 2, BVS

<400> 16

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asn Gly Glu Pro His Met Ile Val
115 120 125

Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu Phe Lys Thr Glu Val Gly
130 135 140

Val Asn Met Cys Thr Leu Met Ala Met Asp Leu Gly Glu Leu Cys Glu
145 150 155 160

Asp Thr Ile Thr Tyr Glu Cys Pro Leu Leu Arg Gln Asn Glu Pro Glu
165 170 175

Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser Thr Trp Val Thr Tyr Gly
180 185 190

Thr Cys Thr Thr Met Gly Glu His Arg Arg Glu Lys Arg Ser Val Ala
195 200 205

Leu Val Pro His Val Gly Met Gly Leu Glu Thr Arg Thr Glu Thr Trp
210 215 220

Met Ser Ser Glu Gly Ala Trp Lys His Val Gln Arg Ile Glu Thr Trp
225 230 235 240

Ile Leu Arg His Pro Gly Phe Thr Met Met Ala Ala Ile Leu Ala Tyr
245 250 255

Thr Ile Gly Thr Thr His Phe Gln Arg Ala Leu Ile Phe Ile Leu Leu
260 265 270

Thr Ala Val Thr Pro Ser Met Thr Met Arg Cys Ile Gly Met Ser Asn
275 280 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ser Trp Val Asp Ile Val
290 295 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305 310 315 320

Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala Lys Gln Pro Ala Thr Leu
325 330 335

Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr Asn Thr Thr Thr Glu Ser
340 345 350

Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu Asn Glu Glu Gln Asp Lys
355 360 365

Arg Phe Val Cys Lys His Ser Met Val Asp Arg Gly Trp Gly Asn Gly
370 375 380

Cys Gly Leu Phe Gly Lys Gly Gly Ile Val Thr Cys Ala Met Phe Arg
385 390 395 400

Cys Lys Lys Asn Met Glu Gly Lys Val Val Gln Pro Glu Asn Leu Glu
405 410 415

Tyr Thr Ile Val Ile Thr Pro His Ser Gly Glu Glu His Ala Val Gly
420 425 430

Asn Asp Thr Gly Lys His Gly Lys Glu Ile Lys Ile Thr Pro Gln Ser
435 440 445

Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr Gly Thr Val Thr Met Glu
450 455 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Gln
465 470 475 480

Met Glu Asn Lys Ala Trp Leu Val His Arg Gln Trp Phe Leu Asp Leu
485 490 495

Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr Gln Gly Ser Asn Trp Ile
500 505 510

Gln Lys Glu Thr Leu Val Thr Phe Lys Asn Pro His Ala Lys Lys Gln
515 520 525

Asp Val Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu
530 535 540

Thr Gly Ala Thr Glu Ile Gln Met Ser Ser Gly Asn Leu Leu Phe Thr
545 550 555 560

Gly His Leu Lys Cys Arg Leu Arg Met Asp Lys Leu Gln Leu Lys Gly
565 570 575

Met Ser Tyr Ser Met Cys Thr Gly Lys Phe Lys Val Val Lys Glu Ile
580 585 590

Ala Glu Thr Gln His Gly Thr Ile Val Ile Arg Val Gln Tyr Glu Gly
595 600 605

Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu Ile Met Asp Leu Glu Lys
610 615 620

Arg His Val Leu Gly Arg Leu Ile Thr Val Asn Pro Ile Val Thr Glu
625 630 635 640

Lys Asp Ser Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln Leu Lys Leu Asn Trp Phe
660 665 670

Lys Lys Gly Ser Ser Ile Gly Gln Met Phe Glu Thr Thr Met Arg Gly
675 680 685

```
Ala Lys Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
    690                 695                 700

Leu Gly Gly Val Phe Thr Ser Ile Gly Lys Ala Leu His Gln Val Phe
705                 710                 715                 720

Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly Val Ser Trp Thr Met Lys
                725                 730                 735

Ile Leu Ile Gly Val Ile Ile Thr Trp Ile Gly Met Asn Ser Arg Ser
            740                 745                 750

Thr Ser Leu Ser Val Thr Leu Val Leu Val Gly Ile Val Thr Leu Tyr
        755                 760                 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
    770                 775                 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785                 790                 795                 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
                805                 810                 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
            820                 825                 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
        835                 840                 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
    850                 855                 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865                 870                 875                 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
                885                 890                 895
```

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
915 920 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
930 935 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945 950 955 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
1025 1030 1035

Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp
1040 1045 1050

Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn
1055 1060 1065

Arg Gly Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile
1070 1075 1080

Thr Glu Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr
1085 1090 1095

```
Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                 1105                 1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                 1120                 1125

His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                 1135                 1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                 1150                 1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                 1165                 1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                 1180                 1185

Ala Thr  Met Thr Asp Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                 1195                 1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                 1210                 1215

Leu Leu  Arg Lys Leu Thr Ser  Lys Glu Leu Met Met  Thr Thr Ile
    1220                 1225                 1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Ile Pro Glu  Thr Ile Leu
    1235                 1240                 1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                 1255                 1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                 1270                 1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                 1285                 1290
```

```
Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                 1300                 1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                 1315                 1320

Thr Ile  Lys Gly Leu Asn Pro  Thr Ala Ile Phe Leu  Thr Thr Leu
    1325                 1330                 1335

Ser Arg  Thr Ser Lys Lys Arg  Ser Trp Pro Leu Asn  Glu Ala Ile
    1340                 1345                 1350

Met Ala  Val Gly Met Val Ser  Ile Leu Ala Ser Ser  Leu Leu Lys
    1355                 1360                 1365

Asn Asp  Ile Pro Met Thr Gly  Pro Leu Val Ala Gly  Gly Leu Leu
    1370                 1375                 1380

Thr Val  Cys Tyr Val Leu Thr  Gly Arg Ser Ala Asp  Leu Glu Leu
    1385                 1390                 1395

Glu Arg  Ala Ala Asp Val Lys  Trp Glu Asp Gln Ala  Glu Ile Ser
    1400                 1405                 1410

Gly Ser  Ser Pro Ile Leu Ser  Ile Thr Ile Ser Glu  Asp Gly Ser
    1415                 1420                 1425

Met Ser  Ile Lys Asn Glu Glu  Glu Glu Gln Thr Leu  Thr Ile Leu
    1430                 1435                 1440

Ile Arg  Thr Gly Leu Leu Val  Ile Ser Gly Leu Phe  Pro Val Ser
    1445                 1450                 1455

Ile Pro  Ile Thr Ala Ala Ala  Trp Tyr Leu Trp Glu  Val Lys Lys
    1460                 1465                 1470

Gln Arg  Ala Gly Val Leu Trp  Asp Val Pro Ser Pro  Pro Pro Met
    1475                 1480                 1485
```

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys
1490 1495 1500

Gly Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu
1505 1510 1515

Gly Thr Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu
1520 1525 1530

Met His Lys Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys
1535 1540 1545

Lys Asp Leu Ile Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu
1550 1555 1560

Trp Lys Glu Gly Glu Glu Val Gln Val Leu Ala Leu Glu Pro Gly
1565 1570 1575

Lys Asn Pro Arg Ala Val Gln Thr Lys Pro Gly Leu Phe Lys Thr
1580 1585 1590

Asn Ala Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser Pro Gly
1595 1600 1605

Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly
1610 1615 1620

Leu Tyr Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser
1625 1630 1635

Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile
1640 1645 1650

Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu
1655 1660 1665

His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val
1670 1675 1680

```
Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                 1690                 1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                 1705                 1710

Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                 1720                 1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                 1735                 1740

Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                 1750                 1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                 1765                 1770

Tyr Ile  Ser Thr Arg Val Glu  Met Gly Glu Ala Ala  Gly Ile Phe
    1775                 1780                 1785

Met Thr  Ala Thr Pro Pro Gly  Ser Arg Asp Pro Phe  Pro Gln Ser
    1790                 1795                 1800

Asn Ala  Pro Ile Ile Asp Glu  Glu Arg Glu Ile Pro  Glu Arg Ser
    1805                 1810                 1815

Trp Asn  Ser Gly His Glu Trp  Val Thr Asp Phe Lys  Gly Lys Thr
    1820                 1825                 1830

Val Trp  Phe Val Pro Ser Ile  Lys Ala Gly Asn Asp  Ile Ala Ala
    1835                 1840                 1845

Cys Leu  Arg Lys Asn Gly Lys  Lys Val Ile Gln Leu  Ser Arg Lys
    1850                 1855                 1860

Thr Phe  Asp Ser Glu Tyr Val  Lys Thr Arg Thr Asn  Asp Trp Asp
    1865                 1870                 1875
```

Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys
1880 1885 1890

Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile
1895 1900 1905

Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro
1910 1915 1920

Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
1925 1930 1935

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
1955 1960 1965

Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met
1970 1975 1980

Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr
1985 1990 1995

Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg
2000 2005 2010

Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu
2015 2020 2025

Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys
2030 2035 2040

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met
2090 2095 2100

Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu
2105 2110 2115

Asp Asn Leu Ala Val Leu His Thr Ala Glu Pro Gly Gly Arg Ala
2120 2125 2130

Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu
2135 2140 2145

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Phe Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg
2525 2530 2535

Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu
2540 2545 2550

Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val
2555 2560 2565

Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly
2570 2575 2580

Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly
2585 2590 2595

Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn
2600 2605 2610

Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro
2615 2620 2625

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp
2810 2815 2820

Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr
2825 2830 2835

Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
2840 2845 2850

Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr
2855 2860 2865

Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg
2870 2875 2880

Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
2885 2890 2895

Ala Leu Gly Ala Val Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala
2900 2905 2910

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys
2945 2950 2955

Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg
2960 2965 2970

Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp
2975 2980 2985

Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu
2990 2995 3000

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
3080 3085 3090

Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala
3095 3100 3105

Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile
3110 3115 3120

Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
3125 3130 3135

Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly
3140 3145 3150

Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
3155 3160 3165

Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln
3170 3175 3180

Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro
3185 3190 3195

Phe Cys Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg
3200 3205 3210

Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg
3215 3220 3225

Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala
3230 3235 3240

Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe
3245 3250 3255

His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
3260 3265 3270

Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
3275 3280 3285

His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
3290 3295 3300

Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 17
<211> 10717
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 3, Tiền chủng gốc

<400> 17
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgatg gagagccgcg catgattgtg 480

gggaagaatg aaagaggaaa atccctactt ttcaagacag cctctggaat caacatgtgc 540

acactcatag ctatggatct gggagagatg tgtgatgaca cggtcactta caaatgcccc 600

cacattaccg aagtggagcc tgaagacatt gactgctggt gcaaccttac atcgacatgg 660

gtgacttatg gaacatgcaa tcaagctgga gagcatagac gcgataagag atcagtggcg 720

ttagctcccc atgttggcat gggactggac acacgcactc aaacctggat gtcggctgaa 780

ggagcttgga gacaagtcga gaaggtagag acatgggccc ttaggcaccc agggtttacc 840

atactagccc tatttcttgc ccattacata ggcacttcct tgacccagaa agtggttatt 900

tttatactat taatgctggt taccccatcc atgacaatga gatgtgtagg agtaggaaac 960

agagattttg tggaaggcct atcgggagct acgtgggttg acgtggtgct cgagcacggt 1020

gggtgtgtga ctaccatggc taagaacaag cccacgctgg acatagagct tcagaagacc 1080

gaggccaccc aactggcgac cctaaggaag ctatgcattg agggaaaaat taccaacata 1140

acaaccgact caagatgtcc cacccaaggg gaagcgattt tacctgagga gcaggaccag 1200

aactacgtgt gtaagcatac atacgtggac agaggctggg gaaacggttg tggtttgttt 1260

ggcaagggaa gcttggtgac atgcgcgaaa tttcaatgtt tagaatcaat agagggaaaa 1320

gtggtgcaac atgagaacct caaatacacc gtcatcatca cagtgcacac aggagaccaa 1380

caccaggtgg gaaatgaaac gcagggagtc acggctgaga taacacccca ggcatcaacc 1440

gctgaagcca ttttacctga atatggaacc ctcgggctag aatgctcacc acggacaggt 1500

ttggatttca atgaaatgat ctcattgaca atgaagaaca aagcatggat ggtacataga 1560

caatggttct ttgacttacc cctaccatgg acatcaggag cttcagcaga aacaccaact 1620

tggaacagga aagagcttct tgtgacattt aaaaatgcac atgcaaaaaa gcaagaagta 1680

gttgttcttg gatcacaaga gggagcaatg catacagcac tgacaggagc tacagagatc 1740

caaacctcag gaggcacaag tatctttgcg gggcacttaa aatgtagact caagatggac 1800

aaattggaac tcaaggggat gagctatgca atgtgcttga gtagctttgt gttgaagaaa 1860

gaagtctccg aaacgcagca tgggacaata ctcattaagg ttgagtacaa aggggaagat 1920

gcaccctgca agattccttt ctccacggag gatggacaag gaaaagctct caatggcaga 1980

ctgatcacag ccaatccagt ggtgaccaag aaggaggagc ctgtcaacat tgaggctgaa 2040

cctccttttg gagaaagtaa catagtaatt ggaattggag acaaagccct gaaaatcaac 2100

tggtacaaga agggaagctc gattgggaag atgttcgagg ccactgccag aggtgcaagg 2160

cgcatggcca tcttgggaga cacagcctgg gactttggat cagtgggtgg tgttttgaat 2220

tcattaggga aaatggtcca ccaaatattt gggagtgctt acacagccct atttggtgga 2280

gtctcctgga tgatgaaaat tggaataggt gtcctcttaa cctggatagg gttgaactca 2340

aaaaatactt ctatgtcatt ttcatgcatc gcggccggca ttgtgacact gtatttggga 2400

gtcatggtgc aggccgatag tggttgcgtt gtgagctgga aaaacaaaga actgaaatgt 2460

ggcagtggga ttttcatcac agacaacgtg cacacatgga cagaacaata caagttccaa 2520

ccagaatccc cttcaaaact agcttcagct atccagaaag cccatgaaga ggacatttgt 2580

ggaatccgct cagtaacaag actggagaat ctgatgtgga aacaaataac accagaattg 2640

aatcacattc tatcagaaaa tgaggtgaag ttaactatta tgacaggaga catcaaagga 2700

atcatgcagg caggaaaacg atctctgcgg cctcagccca ctgagctgaa gtattcatgg 2760

aaaacatggg gcaaagcaaa aatgctctct acagagtctc ataaccagac ctttctcatt 2820

gatggccccg aaacagcaga atgccccaac acaaatagag cttggaattc gttggaagtt 2880

gaagactatg gctttggagt attcaccacc aatatatggc taaaattgaa agaaaaacag 2940

gatgtattct gcgactcaaa actcatgtca gcggccataa aagacaacag agccgtccat 3000

gccgatatgg gttattggat agaaagtgca ctcaatgaca catggaagat agagaaagcc 3060

tctttcattg aagttaaaaa ctgccactgg ccaaaatcac acaccctctg gagcaatgga 3120

gtgctagaaa gtgagatgat aattccaaag aatctcgctg gaccagtgtc tcaacacaac 3180

tatagaccag gctaccatac acaaataaca ggaccatggc atctaggtaa gcttgagatg 3240

gactttgatt tctgtgatgg aacaacagtg gtagtgactg aggactgcgg aaatagagga 3300

ccctctttga gaacaaccac tgcctctgga aaactcataa cagaatggtg ctgccgatct 3360

tgcacattac caccgctaag atacagaggt gaggatgggt gctggtacgg gatggaaatc 3420

agaccattga aggagaaaga agagaatttg gtcaactcct tggtcacagc tggacatggg 3480

caggtcgaca acttttcact aggagtcttg ggaatggcat tgttcctgga ggaaatgctt 3540

aggacccgag taggaacgaa acatgcaata ctactagttg cagtttcttt tgtgacattg 3600

atcacaggga acatgtcctt tagagacctg ggaagagtga tggttatggt aggcgccact 3660

atgacggatg acataggtat gggcgtgact tatcttgccc tactagcagc cttcaaagtc 3720

agaccaactt ttgcagctgg actactcttg agaaagctga cctccaagga attgatgatg 3780

actactatag gaattgtact cctctcccag agcaccatac cagagaccat tcttgagttg 3840

actgatgcgt tagccttagg catgatggtc ctcaaaatgg tgagaaatat ggaaaagtat 3900

caattggcag tgactatcat ggctatcttg tgcgtcccaa acgcagtgat attacaaaac 3960

gcatggaaag tgagttgcac aatattggca gtggtgtccg tttccccact gttcttaaca 4020

tcctcacagc aaaaaacaga ttggatacca ttagcattga cgatcaaagg tctcaatcca 4080

acagctattt ttctaacaac cctctcaaga accagcaaga aaaggagctg gccattaaat 4140

gaggctatca tggcagtcgg gatggtgagc attttagcca gttctctcct aaaaaatgat 4200

attcccatga caggaccatt agtggctgga gggctcctca ctgtgtgcta cgtgctcact 4260

ggacgatcgg ccgatttgga actggagaga gcagccgatg tcaaatggga agaccaggca 4320

gagatatcag gaagcagtcc aatcctgtca ataacaatat cagaagatgg tagcatgtcg 4380

ataaaaaatg aagaggaaga acaaacactg accatactca ttagaacagg attgctggtg 4440

atctcaggac tttttcctgt atcaatacca atcacggcag cagcatggta cctgtgggaa 4500

gtgaagaaac aacgggccgg agtattgtgg gatgttcctt caccccccacc catgggaaag 4560

gctgaactgg aagatggagc ctatagaatt aagcaaaaag ggattcttgg atattcccag 4620

atcggagccg gagtttacaa agaaggaaca ttccatacaa tgtggcatgt cacacgtggc 4680

gctgttctaa tgcataaagg aaagaggatt gaaccatcat gggcggacgt caagaaagac 4740

ctaatatcat atggaggagg ctggaagtta gaaggagaat ggaaggaagg agaagaagtc 4800

caggtattgg cactggagcc tggaaaaaat ccaagagccg tccaaacgaa acctggtctt 4860

ttcaaaacca acgccggaac aataggtgct gtatctctgg acttttctcc tggaacgtca 4920

ggatctccaa ttatcgacaa aaaaggaaaa gttgtgggtc tttatggtaa tggtgttgtt 4980

acaaggagtg gagcatatgt gagtgctata gcccagactg aaaaaagcat tgaagacaac 5040

ccagagatcg aagatgacat tttccgaaag agaagactga ccatcatgga cctccaccca 5100

ggagcgggaa agacgaagag ataccttccg gccatagtca gagaagctat aaaacggggt 5160

ttgagaacat taatcttggc ccccactaga gttgtggcag ctgaaatgga ggaagccctt 5220

agaggacttc caataagata ccagacccca gccatcagag ctgtgcacac cgggcgggag 5280

attgtggacc taatgtgtca tgccacattt accatgaggc tgctatcacc agttagagtg 5340

ccaaactaca acctgattat catggacgaa gcccatttca cagacccagc aagtatagca 5400

gctagaggat acatctcaac tcgagtggag atgggtgagg cagctgggat ttttatgaca 5460

gccactcccc cgggaagcag agacccattt cctcagagca atgcaccaat catagatgaa 5520

gaaagagaaa tccctgaacg ctcgtggaat tccggacatg aatgggtcac ggattttaaa 5580

gggaagactg tttggttcgt tccaagtata aaagcaggaa atgatatagc agcttgcctg 5640

aggaaaaatg gaaagaaagt gatacaactc agtaggaaga cctttgattc tgagtatgtc 5700

aagactagaa ccaatgattg ggacttcgtg gttacaactg acatttcaga aatgggtgcc 5760

aatttcaagg ctgagagggt tatagacccc agacgctgca tgaaaccagt catactaaca 5820

gatggtgaag agcgggtgat tctggcagga cctatgccag tgacccactc tagtgcagca 5880

caaagaagag ggagaatagg aagaaatcca aaaaatgaga atgaccagta catatacatg 5940

ggggaacctc tggaaaatga tgaagactgt gcacactgga aagaagctaa aatgctccta 6000

gataacatca acacgccaga aggaatcatt cctagcatgt tcgaaccaga gcgtgaaaag 6060

gtggatgcca ttgatggcga ataccgcttg agaggagaag caaggaaaac ctttgtagac 6120

ttaatgagaa gaggagacct accagtctgg ttggcctaca gagtggcagc tgaaggcatc 6180

aactacgcag acagaaggtg gtgttttgat ggagtcaaga acaaccaaat cctagaagaa 6240

aacgtggaag ttgaaatctg gacaaaagaa ggggaaagga agaaattgaa acccagatgg 6300

ttggatgcta ggatctattc tgacccactg gcgctaaaag aatttaagga atttgcagcc 6360

ggaagaaagt ctctgaccct gaacctaatc acagaaatgg gtaggctccc aaccttcatg 6420

actcagaagg caagagacgc actggacaac ttagcagtgc tgcacacggc tgaggcaggt 6480

ggaagggcgt acaaccatgc tctcagtgaa ctgccggaga ccctggagac attgctttta 6540

ctgacacttc tggctacagt cacgggaggg atctttttat tcttgatgag cgcaaggggc 6600

atagggaaga tgaccctggg aatgtgctgc ataatcacgg ctagcatcct cctatggtac 6660

gcacaaatac agccacactg gatagcagct tcaataatac tggagttttt tctcatagtt 6720

ttgcttattc cagaacctga aaaacagaga acaccccaag acaaccaact gacctacgtt 6780

gtcatagcca tcctcacagt ggtggccgca accatggcaa acgagatggg tttcctagaa 6840

aaaacgaaga aagatctcgg attgggaagc attgcaaccc agcaacccga gagcaacatc 6900

ctggacatag atctacgtcc tgcatcagca tggacgctgt atgccgtggc cacaacattt 6960

gttacaccaa tgttgagaca tagcattgaa aattcctcag tgaatgtgtc cctaacagct 7020

atagccaacc aagccacagt gttaatgggt ctcgggaaag gatggccatt gtcaaagatg 7080

gacatcggag ttccccttct cgccattgga tgctactcac aagtcaaccc cataactctc 7140

acagcagctc ttttcttatt ggtagcacat tatgccatca tagggccagg actccaagca 7200

aaagcaacca gagaagctca gaaaagagca gcggcgggca tcatgaaaaa cccaactgtc 7260

gatggaataa cagtgattga cctagatcca ataccttatg atccaaagtt tgaaaagcag 7320

ttgggacaag taatgctcct agtcctctgc gtgactcaag tattgatgat gaggactaca 7380

tgggctctgt gtgaggcttt aaccttagct accgggccca tctccacatt gtgggaagga 7440

aatccaggga ggttttggaa cactaccatt gcggtgtcaa tggctaacat ttttagaggg 7500

agttacttgg ccggagctgg acttctcttt tctattatga agaacacaac caacacaaga 7560

aggggaactg gcaacatagg agagacgctt ggagagaaat ggaaaagccg attgaacgcg 7620

ttgggaaaaa gtgaattcca gatctacaag aaaagtggaa tccaggaagt ggatagaacc 7680

ttagcaaaag aaggcattaa aagaggagaa acggaccatc acgctgtgtc gcgaggctca 7740

gcaaaactga gatggttcgt tgagagaaac atggtcacac cagaagggaa agtagtggac 7800

ctcggttgtg gcagaggagg ctggtcatac tattgtggag gactaaagaa tgtaagagaa 7860

gtcaaaggcc taacaaaagg aggaccagga cacgaagaac ccatccccat gtcaacatat 7920

gggtggaatc tagtgcgtct tcaaagtgga gttgacgttt tcttcatccc gccagaaaag 7980

tgtgacacat tattgtgtga cataggggag tcatcaccaa atcccacagt ggaagcagga 8040

cgaacactca gagtccttaa cttagtagaa aattggttga acaacaacac tcaattttgc 8100

ataaaggttc tcaacccata tatgccctca gtcatagaaa aaatggaagc actacaaagg 8160

aaatatggag gagccttagt gaggaatcca ctctcacgaa actccacaca tgagatgtac 8220

tgggtatcca atgcttccgg gaacatagtg tcatcagtga acatgatttc aaggatgttg 8280

atcaacagat ttacaatgag atacaagaaa gccacttacg agccggatgt tgacctcgga 8340

agcggaaccc gtaacatcgg gattgaaagt gagataccaa acctagatat aattgggaaa 8400

agaatagaaa aaataaagca agagcatgaa acatcatggc actatgacca agaccaccca 8460

tacaaaacgt gggcatacca tggtagctat gaaacaaaac agactggatc agcatcatcc 8520

atggtcaacg gagtggtcag gctgctgaca aaaccttggg acgtcgtccc catggtgaca 8580

cagatggcaa tgacagacac gactccattt ggacaacagc gcgtttttaa agagaaagtg 8640

gacacgagaa cccaagaacc gaaagaaggc acgaagaaac taatgaaaat aacagcagag 8700

tggctttgga aagaattagg gaagaaaaag acacccagga tgtgcaccag agaagaattc 8760

acaagaaagg tgagaagcaa tgcagccttg ggggccatat tcactgatga gaacaagtgg 8820

aagtcggcac gtgaggctgt tgaagatagt aggttttggg agctggttga caaggaaagg 8880

aatctccatc ttgaaggaaa gtgtgaaaca tgtgtgtaca acatgatggg aaaaagagag 8940

aagaagctag gggaattcgg caaggcaaaa ggcagcagag ccatatggta catgtggctt 9000

ggagcacgct tcttagagtt tgaagcccta ggattcttaa atgaagatca ctggttctcc 9060

agagagaact ccctgagtgg agtggaagga gaagggctgc acaagctagg ttacattcta 9120

agagacgtga gcaagaaaga gggaggagca atgtatgccg atgacaccgc aggatgggat 9180

acaagaatca cactagaaga cctaaaaaat gaagaaatgg taacaaacca catggaagga 9240

gaacacaaga aactagccga ggccattttc aaactaacgt accaaaacaa ggtggtgcgt 9300

gtgcaaagac caacaccaag aggcacagta atggacatca tatcgagaag agaccaaaga 9360

ggtagtggac aagttggcac ctatggactc aatactttca ccaatatgga agcccaacta 9420

atcagacaga tggagggaga aggagtcttt aaaagcattc agcacctaac aatcacagaa 9480

gaaatcgctg tgcaaaactg gttagcaaga gtggggcgcg aaaggttatc aagaatggcc 9540

atcagtggag atgattgtgt tgtgaaacct ttagatgaca ggttcgcaag cgctttaaca 9600

gctctaaatg acatgggaaa gattaggaaa gacatacaac aatgggaacc ttcaagagga 9660

tggaatgatt ggacacaagt gcccttctgt tcacaccatt tccatgagtt aatcatgaaa 9720

gacggtcgcg tactcgttgt tccatgtaga aaccaagatg aactgattgg cagagcccga 9780

atctcccaag gagcagggtg gtctttgcgg gagacggcct gtttggggaa gtcttacgcc 9840

caaatgtgga gcttgatgta cttccacaga cgcgacctca ggctggcggc aaatgctatt 9900

tgctcggcag taccatcaca ttgggttcca acaagtcgaa caacctggtc catacatgct 9960

aaacatgaat ggatgacaac ggaagacatg ctgacagtct ggaacagggt gtggattcaa 10020

gaaaacccat ggatggaaga caaaactcca gtggaatcat gggaggaaat cccatacttg 10080

gggaaaagag aagaccaatg gtgcggctca ttgattgggt taacaagcag ggccacctgg 10140

gcaaagaaca tccaagcagc aataaatcaa gttagatccc ttataggcaa tgaagaatac 10200

acagattaca tgccatccat gaaaagattc agaagagaag aggaagaagc aggagttctg 10260

tggtagaaag caaaactaac atgaaacaag gctagaagtc aggtcggatt aagccatagt 10320

acggaaaaaa ctatgctacc tgtgagcccc gtccaaggac gttaaaagaa gtcaggccat 10380

cataaatgcc atagcttgag taaactatgc agcctgtagc tccacctgag aaggtgtaaa 10440

aaatccggga ggccacaaac catggaagct gtacgcatgg cgtagtggac tagcggttag 10500

aggagacccc tcccttacaa atcgcagcaa caatgggggc ccaaggcgag atgaagctgt 10560

agtctcgctg gaaggactag aggttagagg agacccccccc gaaacaaaaa acagcatatt 10620

gacgctggga aagaccagag atcctgctgt ctcctcagca tcattccagg cacagaacgc 10680

cagaaaatgg aatggtgctg ttgaatcaac aggttct 10717

<210> 18
<211> 3389
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 3, Tiền chủng gốc

<400> 18

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Arg Met Ile Val
115 120 125

Gly Lys Asn Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ala Ser Gly
130 135 140

Ile Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Asp
145 150 155 160

Asp Thr Val Thr Tyr Lys Cys Pro His Ile Thr Glu Val Glu Pro Glu
165 170 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Thr Tyr Gly
180 185 190

Thr Cys Asn Gln Ala Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
195 200 205

```
Leu Ala Pro His Val Gly Met Gly Leu Asp Thr Arg Thr Gln Thr Trp
    210                 215                 220

Met Ser Ala Glu Gly Ala Trp Arg Gln Val Glu Lys Val Glu Thr Trp
225                 230                 235                 240

Ala Leu Arg His Pro Gly Phe Thr Ile Leu Ala Leu Phe Leu Ala His
                245                 250                 255

Tyr Ile Gly Thr Ser Leu Thr Gln Lys Val Val Ile Phe Ile Leu Leu
            260                 265                 270

Met Leu Val Thr Pro Ser Met Thr Met Arg Cys Val Gly Val Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
    290                 295                 300

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Ile Glu Leu Gln Lys Thr Glu Ala Thr Gln Leu Ala Thr Leu
                325                 330                 335

Arg Lys Leu Cys Ile Glu Gly Lys Ile Thr Asn Ile Thr Thr Asp Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Ala Ile Leu Pro Glu Glu Gln Asp Gln
        355                 360                 365

Asn Tyr Val Cys Lys His Thr Tyr Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Val Thr Cys Ala Lys Phe Gln
385                 390                 395                 400

Cys Leu Glu Ser Ile Glu Gly Lys Val Val Gln His Glu Asn Leu Lys
                405                 410                 415
```

```
Tyr Thr Val Ile Ile Thr Val His Thr Gly Asp Gln His Gln Val Gly
            420                 425                 430

Asn Glu Thr Gln Gly Val Thr Ala Glu Ile Thr Pro Gln Ala Ser Thr
        435                 440                 445

Ala Glu Ala Ile Leu Pro Glu Tyr Gly Thr Leu Gly Leu Glu Cys Ser
    450                 455                 460

Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Ile Ser Leu Thr Met Lys
465                 470                 475                 480

Asn Lys Ala Trp Met Val His Arg Gln Trp Phe Phe Asp Leu Pro Leu
                485                 490                 495

Pro Trp Thr Ser Gly Ala Ser Ala Glu Thr Pro Thr Trp Asn Arg Lys
            500                 505                 510

Glu Leu Leu Val Thr Phe Lys Asn Ala His Ala Lys Lys Gln Glu Val
        515                 520                 525

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

Ala Thr Glu Ile Gln Thr Ser Gly Gly Thr Ser Ile Phe Ala Gly His
545                 550                 555                 560

Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Glu Leu Lys Gly Met Ser
                565                 570                 575

Tyr Ala Met Cys Leu Ser Ser Phe Val Leu Lys Lys Glu Val Ser Glu
            580                 585                 590

Thr Gln His Gly Thr Ile Leu Ile Lys Val Glu Tyr Lys Gly Glu Asp
        595                 600                 605

Ala Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp Gly Gln Gly Lys Ala
    610                 615                 620
```

Leu Asn Gly Arg Leu Ile Thr Ala Asn Pro Val Val Thr Lys Lys Glu
625 630 635 640

Glu Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser Asn Ile
645 650 655

Val Ile Gly Ile Gly Asp Lys Ala Leu Lys Ile Asn Trp Tyr Lys Lys
660 665 670

Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly Ala Arg
675 680 685

Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly
690 695 700

Gly Val Leu Asn Ser Leu Gly Lys Met Val His Gln Ile Phe Gly Ser
705 710 715 720

Ala Tyr Thr Ala Leu Phe Gly Gly Val Ser Trp Met Met Lys Ile Gly
725 730 735

Ile Gly Val Leu Leu Thr Trp Ile Gly Leu Asn Ser Lys Asn Thr Ser
740 745 750

Met Ser Phe Ser Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr Leu Gly
755 760 765

Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys Asn Lys
770 775 780

Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val His Thr
785 790 795 800

Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys Leu Ala
805 810 815

Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile Arg Ser
820 825 830

```
Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro Glu Leu
        835                 840                 845

Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met Thr Gly
    850                 855                 860

Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg Pro Gln
865                 870                 875                 880

Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala Lys Met
                885                 890                 895

Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly Pro Glu
            900                 905                 910

Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu Glu Val
        915                 920                 925

Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu Lys Leu
    930                 935                 940

Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser Ala Ala
945                 950                 955                 960

Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp Ile Glu
                965                 970                 975

Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe Ile Glu
            980                 985                 990

Val Lys Asn Cys His Trp Pro Lys  Ser His Thr Leu Trp  Ser Asn Gly
        995                 1000                1005

Val Leu  Glu Ser Glu Met Ile  Ile Pro Lys Asn Leu  Ala Gly Pro
    1010                1015                1020

Val Ser  Gln His Asn Tyr Arg  Pro Gly Tyr His Thr  Gln Ile Thr
    1025                1030                1035
```

```
Gly Pro  Trp His Leu Gly Lys  Leu Glu Met Asp Phe  Asp Phe Cys
    1040                 1045                 1050

Asp Gly  Thr Thr Val Val Val  Thr Glu Asp Cys Gly  Asn Arg Gly
    1055                 1060                 1065

Pro Ser  Leu Arg Thr Thr Thr  Ala Ser Gly Lys Leu  Ile Thr Glu
    1070                 1075                 1080

Trp Cys  Cys Arg Ser Cys Thr  Leu Pro Pro Leu Arg  Tyr Arg Gly
    1085                 1090                 1095

Glu Asp  Gly Cys Trp Tyr Gly  Met Glu Ile Arg Pro  Leu Lys Glu
    1100                 1105                 1110

Lys Glu  Glu Asn Leu Val Asn  Ser Leu Val Thr Ala  Gly His Gly
    1115                 1120                 1125

Gln Val  Asp Asn Phe Ser Leu  Gly Val Leu Gly Met  Ala Leu Phe
    1130                 1135                 1140

Leu Glu  Glu Met Leu Arg Thr  Arg Val Gly Thr Lys  His Ala Ile
    1145                 1150                 1155

Leu Leu  Val Ala Val Ser Phe  Val Thr Leu Ile Thr  Gly Asn Met
    1160                 1165                 1170

Ser Phe  Arg Asp Leu Gly Arg  Val Met Val Met Val  Gly Ala Thr
    1175                 1180                 1185

Met Thr  Asp Asp Ile Gly Met  Gly Val Thr Tyr Leu  Ala Leu Leu
    1190                 1195                 1200

Ala Ala  Phe Lys Val Arg Pro  Thr Phe Ala Ala Gly  Leu Leu Leu
    1205                 1210                 1215

Arg Lys  Leu Thr Ser Lys Glu  Leu Met Met Thr Thr  Ile Gly Ile
    1220                 1225                 1230
```

Val Leu Leu Ser Gln Ser Thr Ile Pro Glu Thr Ile Leu Glu Leu
1235 1240 1245

Thr Asp Ala Leu Ala Leu Gly Met Met Val Leu Lys Met Val Arg
1250 1255 1260

Asn Met Glu Lys Tyr Gln Leu Ala Val Thr Ile Met Ala Ile Leu
1265 1270 1275

Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys Val Ser
1280 1285 1290

Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe Leu Thr
1295 1300 1305

Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile
1310 1315 1320

Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg
1325 1330 1335

Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala
1340 1345 1350

Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp
1355 1360 1365

Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu Thr Val
1370 1375 1380

Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu Glu Arg
1385 1390 1395

Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser Gly Ser
1400 1405 1410

Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser Met Ser
1415 1420 1425

Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu Ile Arg
1430 1435 1440

Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser Ile Pro
1445 1450 1455

Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys Gln Arg
1460 1465 1470

Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met Gly Lys
1475 1480 1485

Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly Ile
1490 1495 1500

Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr
1505 1510 1515

Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu Met His
1520 1525 1530

Lys Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys Lys Asp
1535 1540 1545

Leu Ile Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu Trp Lys
1550 1555 1560

Glu Gly Glu Glu Val Gln Val Leu Ala Leu Glu Pro Gly Lys Asn
1565 1570 1575

Pro Arg Ala Val Gln Thr Lys Pro Gly Leu Phe Lys Thr Asn Ala
1580 1585 1590

Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser Pro Gly Thr Ser
1595 1600 1605

Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly Leu Tyr
1610 1615 1620

Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser Ala Ile
1625 1630 1635

Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile Glu Asp
1640 1645 1650

Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu His Pro
1655 1660 1665

Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val Arg Glu
1670 1675 1680

Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro Thr Arg
1685 1690 1695

Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu Pro Ile
1700 1705 1710

Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly Arg Glu
1715 1720 1725

Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu Leu
1730 1735 1740

Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
1745 1750 1755

Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile
1760 1765 1770

Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr
1775 1780 1785

Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala
1790 1795 1800

Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn
1805 1810 1815

```
Ser Gly  His Glu Trp Val Thr  Asp Phe Lys Gly Lys  Thr Val Trp
    1820                 1825                 1830

Phe Val  Pro Ser Ile Lys Ala  Gly Asn Asp Ile Ala  Ala Cys Leu
    1835                 1840                 1845

Arg Lys  Asn Gly Lys Lys Val  Ile Gln Leu Ser Arg  Lys Thr Phe
    1850                 1855                 1860

Asp Ser  Glu Tyr Val Lys Thr  Arg Thr Asn Asp Trp  Asp Phe Val
    1865                 1870                 1875

Val Thr  Thr Asp Ile Ser Glu  Met Gly Ala Asn Phe  Lys Ala Glu
    1880                 1885                 1890

Arg Val  Ile Asp Pro Arg Arg  Cys Met Lys Pro Val  Ile Leu Thr
    1895                 1900                 1905

Asp Gly  Glu Glu Arg Val Ile  Leu Ala Gly Pro Met  Pro Val Thr
    1910                 1915                 1920

His Ser  Ser Ala Ala Gln Arg  Arg Gly Arg Ile Gly  Arg Asn Pro
    1925                 1930                 1935

Lys Asn  Glu Asn Asp Gln Tyr  Ile Tyr Met Gly Glu  Pro Leu Glu
    1940                 1945                 1950

Asn Asp  Glu Asp Cys Ala His  Trp Lys Glu Ala Lys  Met Leu Leu
    1955                 1960                 1965

Asp Asn  Ile Asn Thr Pro Glu  Gly Ile Ile Pro Ser  Met Phe Glu
    1970                 1975                 1980

Pro Glu  Arg Glu Lys Val Asp  Ala Ile Asp Gly Glu  Tyr Arg Leu
    1985                 1990                 1995

Arg Gly  Glu Ala Arg Lys Thr  Phe Val Asp Leu Met  Arg Arg Gly
    2000                 2005                 2010
```

Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile
2015 2020 2025

Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn
2030 2035 2040

Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu
2045 2050 2055

Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile
2060 2065 2070

Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala
2075 2080 2085

Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met Gly Arg
2090 2095 2100

Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu Asp Asn
2105 2110 2115

Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala Tyr Asn
2120 2125 2130

His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu Leu Leu
2135 2140 2145

Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu Phe Leu
2150 2155 2160

Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met Cys Cys
2165 2170 2175

Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile Gln Pro
2180 2185 2190

His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu Ile Val
2195 2200 2205

Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp Asn
2210 2215 2220

Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val Ala Ala
2225 2230 2235

Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp
2240 2245 2250

Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile
2255 2260 2265

Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala
2270 2275 2280

Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu
2285 2290 2295

Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala
2300 2305 2310

Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser Lys Met
2315 2320 2325

Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser Gln Val
2330 2335 2340

Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val Ala His
2345 2350 2355

Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr Arg Glu
2360 2365 2370

Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro Thr Val
2375 2380 2385

Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr Asp Pro
2390 2395 2400

Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val Leu Cys
2405 2410 2415

Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu Cys Glu
2420 2425 2430

Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly
2435 2440 2445

Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala
2450 2455 2460

Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe
2465 2470 2475

Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn
2480 2485 2490

Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala
2495 2500 2505

Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln
2510 2515 2520

Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu
2525 2530 2535

Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp
2540 2545 2550

Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val Val Asp
2555 2560 2565

Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly Gly Leu
2570 2575 2580

Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly Pro Gly
2585 2590 2595

His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn Leu Val
2600 2605 2610

Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu Lys
2615 2620 2625

Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro Asn Pro
2630 2635 2640

Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu Val Glu
2645 2650 2655

Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val Leu Asn
2660 2665 2670

Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu Gln Arg
2675 2680 2685

Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn Ser
2690 2695 2700

Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val
2705 2710 2715

Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr
2720 2725 2730

Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly
2735 2740 2745

Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu
2750 2755 2760

Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu His Glu
2765 2770 2775

Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr Trp Ala
2780 2785 2790

Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala Ser Ser
2795 2800 2805

Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp Asp Val
2810 2815 2820

Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr Pro Phe
2825 2830 2835

Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg Thr Gln
2840 2845 2850

Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr Ala Glu
2855 2860 2865

Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg Met Cys
2870 2875 2880

Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala Ala Leu
2885 2890 2895

Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala Arg Glu
2900 2905 2910

Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys Glu Arg
2915 2920 2925

Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn Met
2930 2935 2940

Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
2945 2950 2955

Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu
2960 2965 2970

Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser
2975 2980 2985

```
Arg Glu  Asn Ser Leu Ser Gly  Val Glu Gly Glu Gly  Leu His Lys
    2990                 2995                 3000

Leu Gly  Tyr Ile Leu Arg Asp  Val Ser Lys Lys Glu  Gly Gly Ala
    3005                 3010                 3015

Met Tyr  Ala Asp Asp Thr Ala  Gly Trp Asp Thr Arg  Ile Thr Leu
    3020                 3025                 3030

Glu Asp  Leu Lys Asn Glu Glu  Met Val Thr Asn His  Met Glu Gly
    3035                 3040                 3045

Glu His  Lys Lys Leu Ala Glu  Ala Ile Phe Lys Leu  Thr Tyr Gln
    3050                 3055                 3060

Asn Lys  Val Val Arg Val Gln  Arg Pro Thr Pro Arg  Gly Thr Val
    3065                 3070                 3075

Met Asp  Ile Ile Ser Arg Arg  Asp Gln Arg Gly Ser  Gly Gln Val
    3080                 3085                 3090

Gly Thr  Tyr Gly Leu Asn Thr  Phe Thr Asn Met Glu  Ala Gln Leu
    3095                 3100                 3105

Ile Arg  Gln Met Glu Gly Glu  Gly Val Phe Lys Ser  Ile Gln His
    3110                 3115                 3120

Leu Thr  Ile Thr Glu Glu Ile  Ala Val Gln Asn Trp  Leu Ala Arg
    3125                 3130                 3135

Val Gly  Arg Glu Arg Leu Ser  Arg Met Ala Ile Ser  Gly Asp Asp
    3140                 3145                 3150

Cys Val  Val Lys Pro Leu Asp  Asp Arg Phe Ala Ser  Ala Leu Thr
    3155                 3160                 3165

Ala Leu  Asn Asp Met Gly Lys  Ile Arg Lys Asp Ile  Gln Gln Trp
    3170                 3175                 3180
```

```
Glu Pro  Ser Arg Gly Trp Asn  Asp Trp Thr Gln Val  Pro Phe Cys
    3185                3190                3195

Ser His  His Phe His Glu Leu  Ile Met Lys Asp Gly  Arg Val Leu
    3200                3205                3210

Val Val  Pro Cys Arg Asn Gln  Asp Glu Leu Ile Gly  Arg Ala Arg
    3215                3220                3225

Ile Ser  Gln Gly Ala Gly Trp  Ser Leu Arg Glu Thr  Ala Cys Leu
    3230                3235                3240

Gly Lys  Ser Tyr Ala Gln Met  Trp Ser Leu Met Tyr  Phe His Arg
    3245                3250                3255

Arg Asp  Leu Arg Leu Ala Ala  Asn Ala Ile Cys Ser  Ala Val Pro
    3260                3265                3270

Ser His  Trp Val Pro Thr Ser  Arg Thr Thr Trp Ser  Ile His Ala
    3275                3280                3285

Lys His  Glu Trp Met Thr Thr  Glu Asp Met Leu Thr  Val Trp Asn
    3290                3295                3300

Arg Val  Trp Ile Gln Glu Asn  Pro Trp Met Glu Asp  Lys Thr Pro
    3305                3310                3315

Val Glu  Ser Trp Glu Glu Ile  Pro Tyr Leu Gly Lys  Arg Glu Asp
    3320                3325                3330

Gln Trp  Cys Gly Ser Leu Ile  Gly Leu Thr Ser Arg  Ala Thr Trp
    3335                3340                3345

Ala Lys  Asn Ile Gln Ala Ala  Ile Asn Gln Val Arg  Ser Leu Ile
    3350                3355                3360

Gly Asn  Glu Glu Tyr Thr Asp  Tyr Met Pro Ser Met  Lys Arg Phe
    3365                3370                3375
```

Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385

<210> 19
<211> 10717
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 3, MVS

<400> 19
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgatg gagagccgcg catgattgtg 480

gggaagaatg aaagaggaaa atccctactt ttcaagacag cctctggaat caacatgtgc 540

acactcatag ctatggatct gggagagatg tgtgatgaca cggtcactta caaatgcccc 600

cacattaccg aagtggagcc tgaagacatt gactgctggt gcaaccttac atcgacatgg 660

gtgacttatg gaacatgcaa tcaagctgga gagcatagac gcgataagag atcagtggcg 720

ttagctcccc atgttggcat gggactggac acacgcactc aaacctggat gtcggctgaa 780

ggagcttgga gacaagtcga gaaggtagag acatgggccc ttaggcaccc agggtttacc 840

atactagccc tatttcttgc ccattacata ggcacttcct tgacccagaa agtggttatt 900

tttatactat taatgctggt taccccatcc atgacaatga gatgtgtagg agtaggaaac 960

agagattttg tggaaggcct atcgggagct acgtgggttg acgtggtgct cgagcacggt 1020

gggtgtgtga ctaccatggc taagaacaag cccacgctgg acatagagct tcagaagacc 1080

gaggccaccc aactggcgac cctaaggaag ctatgcattg agggaaaaat taccaacata 1140

acaaccgact caagatgtcc cacccaaggg gaagcgattt tacctgagga gcaggaccag 1200

aactacgtgt gtaagcatac atacgtggac agaggctggg gaaacggttg tggtttgttt 1260

ggcaagggaa gcttggtgac atgcgcgaaa tttcaatgtt tagaatcaat agagggaaaa 1320

gtggtgcaac atgagaacct caaatacacc gtcatcatca cagtgcacac aggagaccaa 1380

caccaggtgg gaaatgaaac gcagggagtc acggctgaga taacacccca ggcatcaacc 1440

gctgaagcca ttttacctga atatggaacc ctcgggctag aatgctcacc acggacaggt 1500

ttggatttca atgaaatgat ctcattgaca atgaagaaca aagcatggat ggtacataga 1560

caatggttct ttgacttacc cctaccatgg acatcaggag cttcagcaga aacaccaact 1620

tggaacagga aagagcttct tgtgacattt aaaaatgcac atgcaaaaaa gcaagaagta 1680

gttgttcttg gatcacaaga gggagcaatg catacagcac tgacaggagc tacagagatc 1740

caaacctcag gaggcacaag tatctttgcg gggcacttaa aatgtagact caagatggac 1800

aaattggaac tcaaggggat gagctatgca atgtgcttga gtagctttgt gttgaagaaa 1860

gaagtctccg aaacgcagca tgggacaata ctcattaagg ttgagtacaa aggggaagat 1920

gcaccctgca agattccttt ctccacggag gatggacaag gaaaagctct caatggcaga 1980

ctgatcacag ccaatccagt ggtgaccaag aaggaggagc ctgtcaacat tgaggctgaa 2040

cctccttttg gagaaagtaa catagtaatt ggaattggag acaaagccct gaaaatcaac 2100

tggtacaaga agggaagctc gattgggaag atgttcgagg ccactgccag aggtgcaagg 2160

cgcatggcca tcttgggaga cacagcctgg gactttggat cagtgggtgg tgttttgaat 2220

tcattaggga aaatggtcca ccaaatattt gggagtgctt acacagccct atttggtgga 2280

gtctcctgga tgatgaaaat tggaataggt gtcctcttaa cctggatagg gttgaactca 2340

aaaaatactt ctatgtcatt ttcatgcatc gcggccggca ttgtgacact gtatttggga 2400

gtcatggtgc aggccgatag tggttgcgtt gtgagctgga aaaacaaaga actgaaatgt 2460

ggcagtggga ttttcatcac agacaacgtg cacacatgga cagaacaata caagttccaa 2520

ccagaatccc cttcaaaact agcttcagct atccagaaag cccatgaaga ggacatttgt 2580

ggaatccgct cagtaacaag actggagaat ctgatgtgga aacaaataac accagaattg 2640

aatcacattc tatcagaaaa tgaggtgaag ttaactatta tgacaggaga catcaaagga 2700

atcatgcagg caggaaaacg atctctgcgg cctcagccca ctgagctgaa gtattcatgg 2760

aaaacatggg gcaaagcaaa aatgctctct acagagtctc ataaccagac ctttctcatt 2820

gatggccccg aaacagcaga atgccccaac acaaatagag cttggaattc gttggaagtt 2880

gaagactatg gctttggagt attcaccacc aatatatggc taaaattgaa agaaaaacag 2940

gatgtattct gcgactcaaa actcatgtca gcggccataa aagacaacag agccgtccat 3000

gccgatatgg gttattggat agaaagtgca ctcaatgaca catggaagat agagaaagcc 3060

tctttcattg aagttaaaaa ctgccactgg ccaaaatcac acaccctctg gagcaatgga 3120

gtgctagaaa gtgagatgat aattccaaag aatctcgctg gaccagtgtc tcaacacaac 3180

tatagaccag gctaccatac acaaataaca ggaccatggc atctaggtaa gcttgagatg 3240

gactttgatt tctgtgatgg aacaacagtg gtagtgactg aggactgcgg aaatagagga 3300

ccctctttga gaacaaccac tgcctctgga aaactcataa cagaatggtg ctgccgatct 3360

tgcacattac caccgctaag atacagaggt gaggatgggt gctggtacgg gatggaaatc 3420

agaccattga aggagaaaga agagaatttg gtcaactcct tggtcacagc tggacatggg 3480

caggtcgaca acttttcact aggagtcttg ggaatggcat tgttcctgga ggaaatgctt 3540

aggacccgag taggaacgaa acatgcaata ctactagttg cagtttcttt tgtgacattg 3600

atcacaggga acatgtcctt tagagacctg ggaagagtga tggttatggt aggcgccact 3660

atgacggatg acataggtat gggcgtgact tatcttgccc tactagcagc cttcaaagtc 3720

agaccaactt ttgcagctgg actactcttg agaaagctga cctccaagga attgatgatg 3780

actactatag gaattgtact cctctcccag agcaccatac cagagaccat tcttgagttg 3840

actgatgcgt tagccttagg catgatggtc ctcaaaatgg tgagaaatat ggaaaagtat 3900

caattggcag tgactatcat ggctatcttg tgcgtcccaa acgcagtgat attacaaaac 3960

gcatggaaag tgagttgcac aatattggca gtggtgtccg tttccccact gttcttaaca 4020

tcctcacagc aaaaaacaga ttggatacca ttagcattga cgatcaaagg tctcaatcca 4080

acagctattt ttctaacaac cctctcaaga accagcaaga aaaggagctg gccattaaat 4140

gaggctatca tggcagtcgg gatggtgagc attttagcca gttctctcct aaaaaatgat 4200

attcccatga caggaccatt agtggctgga gggctcctca ctgtgtgcta cgtgctcact 4260

ggacgatcgg ccgatttgga actggagaga gcagccgatg tcaaatggga agaccaggca 4320

gagatatcag gaagcagtcc aatcctgtca ataacaatat cagaagatgg tagcatgtcg 4380

ataaaaaatg aagaggaaga acaaacactg accatactca ttagaacagg attgctggtg 4440

atctcaggac tttttcctgt atcaatacca atcacggcag cagcatggta cctgtgggaa 4500

gtgaagaaac aacgggccgg agtattgtgg gatgttcctt caccccccacc catgggaaag 4560

gctgaactgg aagatggagc ctatagaatt aagcaaaaag ggattcttgg atattcccag 4620

atcggagccg gagtttacaa agaaggaaca ttccatacaa tgtggcatgt cacacgtggc 4680

gctgttctaa tgcataaagg aaagaggatt gaaccatcat gggcggacgt caagaaagac 4740

ctaatatcat atggaggagg ctggaagtta gaaggagaat ggaaggaagg agaagaagtc 4800

caggtattgg cactggagcc tggaaaaaat ccaagagccg tccaaacgaa acctggtctt 4860

ttcaaaacca acgccggaac aataggtgct gtatctctgg acttttctcc tggaacgtca 4920

ggatctccaa ttatcgacaa aaaaggaaaa gttgtgggtc tttatggtaa tggtgttgtt 4980

acaaggagtg gagcatatgt gagtgctata gcccagactg aaaaaagcat tgaagacaac 5040

ccagagatcg aagatgacat tttccgaaag agaagactga ccatcatgga cctccaccca 5100

ggagcgggaa agacgaagag ataccttccg gccatagtca gagaagctat aaaacggggt 5160

ttgagaacat taatcttggc ccccactaga gttgtggcag ctgaaatgga ggaagccctt 5220

agaggacttc caataagata ccagacccca gccatcagag ctgtgcacac cgggcgggag 5280

attgtggacc taatgtgtca tgccacattt accatgaggc tgctatcacc agttagagtg 5340

ccaaactaca acctgattat catggacgaa gcccatttca cagacccagc aagtatagca 5400

gctagaggat acatctcaac tcgagtggag atgggtgagg cagctgggat ttttatgaca 5460

gccactcccc cgggaagcag agacccattt cctcagagca atgcaccaat catagatgaa 5520

gaaagagaaa tccctgaacg ctcgtggaat tccggacatg aatgggtcac ggattttaaa 5580

gggaagactg tttggttcgt tccaagtata aaagcaggaa atgatatagc agcttgcctg 5640

aggaaaaatg gaaagaaagt gatacaactc agtaggaaga cctttgattc tgagtatgtc 5700

aagactagaa ccaatgattg ggacttcgtg gttacaactg acatttcaga aatgggtgcc 5760

aatttcaagg ctgagagggt tatagacccc agacgctgca tgaaaccagt catactaaca 5820

gatggtgaag agcgggtgat tctggcagga cctatgccag tgacccactc tagtgcagca 5880

caaagaagag ggagaatagg aagaaatcca aaaaatgaga atgaccagta catatacatg 5940

ggggaacctc tggaaaatga tgaagactgt gcacactgga aagaagctaa aatgctccta 6000

gataacatca acacgccaga aggaatcatt cctagcatgt tcgaaccaga gcgtgaaaag 6060

gtggatgcca ttgatggcga ataccgcttg agaggagaag caaggaaaac ctttgtagac 6120

ttaatgagaa gaggagacct accagtctgg ttggcctaca gagtggcagc tgaaggcatc 6180

aactacgcag acagaaggtg gtgttttgat ggagtcaaga acaaccaaat cctagaagaa 6240

aacgtggaag ttgaaatctg gacaaaagaa ggggaaagga agaaattgaa acccagatgg 6300

ttggatgcta ggatctattc tgacccactg gcgctaaaag aatttaagga atttgcagcc 6360

ggaagaaagt ctctgaccct gaacctaatc acagaaatgg gtaggctccc aaccttcatg 6420

actcagaagg caagagacgc actggacaac ttagcagtgc tgcacacggc tgaggcaggt 6480

ggaagggcgt acaaccatgc tctcagtgaa ctgccggaga ccctggagac attgctttta 6540

ctgacacttc tggctacagt cacgggaggg atctttttat tcttgatgag cgcaaggggc 6600

atagggaaga tgaccctggg aatgtgctgc ataatcacgg ctagcatcct cctatggtac 6660

gcacaaatac agccacactg gatagcagct tcaataatac tggagttttt tctcatagtt 6720

ttgcttattc cagaacctga aaaacagaga acaccccaag acaaccaact gacctacgtt 6780

gtcatagcca tcctcacagt ggtggccgca accatggcaa acgagatggg tttcctagaa 6840

aaaacgaaga aagatctcgg attgggaagc attgcaaccc agcaacccga gagcaacatc 6900

ctggacatag atctacgtcc tgcatcagca tggacgctgt atgccgtggc cacaacattt 6960

gttacaccaa tgttgagaca tagcattgaa aattcctcag tgaatgtgtc cctaacagct 7020

atagccaacc aagccacagt gttaatgggt ctcgggaaag gatggccatt gtcaaagatg 7080

gacatcggag ttccccttct cgccattgga tgctactcac aagtcaaccc cataactctc 7140

acagcagctc ttttcttatt ggtagcacat tatgccatca tagggccagg actccaagca 7200

aaagcaacca gagaagctca gaaaagagca gcggcgggca tcatgaaaaa cccaactgtc 7260

gatggaataa cagtgattga cctagatcca ataccttatg atccaaagtt tgaaaagcag 7320

ttgggacaag taatgctcct agtcctctgc gtgactcaag tattgatgat gaggactaca 7380

tgggctctgt gtgaggcttt aaccttagct accgggccca tctccacatt gtgggaagga 7440

aatccaggga ggttttggaa cactaccatt gcggtgtcaa tggctaacat ttttagaggg 7500

agttacttgg ccggagctgg acttctcttt tctattatga agaacacaac caacacaaga 7560

aggggaactg gcaacatagg agagacgctt ggagagaaat ggaaaagccg attgaacgcg 7620

ttgggaaaaa gtgaattcca gatctacaag aaaagtggaa tccaggaagt ggatagaacc 7680

ttagcaaaag aaggcattaa aagaggagaa acggaccatc acgctgtgtc gcgaggctca 7740

gcaaaactga gatggttcgt tgagagaaac atggtcacac cagaagggaa agtagtggac 7800

ctcggttgtg gcagaggagg ctggtcatac tattgtggag gactaaagaa tgtaagagaa 7860

gtcaaaggcc taacaaaagg aggaccagga cacgaagaac ccatccccat gtcaacatat 7920

gggtggaatc tagtgcgtct tcaaagtgga gttgacgttt tcttcatccc gccagaaaag 7980

tgtgacacat tattgtgtga cataggggag tcatcaccaa atcccacagt ggaagcagga 8040

cgaacactca gagtccttaa cttagtagaa aattggttga acaacaacac tcaattttgc 8100

ataaaggttc tcaacccata tatgccctca gtcatagaaa aaatggaagc actacaaagg 8160

aaatatggag gagccttagt gaggaatcca ctctcacgaa actccacaca tgagatgtac 8220

tgggtatcca atgcttccgg gaacatagtg tcatcagtga acatgatttc aaggatgttg 8280

atcaacagat ttacaatgag atacaagaaa gccacttacg agccggatgt tgacctcgga 8340

agcggaaccc gtaacatcgg gattgaaagt gagataccaa acctagatat aattgggaaa 8400

agaatagaaa aaataaagca agagcatgaa acatcatggc actatgacca agaccaccca 8460

tacaaaacgt gggcatacca tggtagctat gaaacaaaac agactggatc agcatcatcc 8520

atggtcaacg gagtggtcag gctgctgaca aaaccttggg acgtcgtccc catggtgaca 8580

cagatggcaa tgacagacac gactccattt ggacaacagc gcgtttttaa agagaaagtg 8640

gacacgagaa cccaagaacc gaaagaaggc acgaagaaac taatgaaaat aacagcagag 8700

tggctttgga aagaattagg gaagaaaaag acacccagga tgtgcaccag agaagaattc 8760

acaagaaagg tgagaagcaa tgcagccttg ggggccatat tcactgatga gaacaagtgg 8820

aagtcggcac gtgaggctgt tgaagatagt aggttttggg agctggttga caaggaaagg 8880

aatctccatc ttgaaggaaa gtgtgaaaca tgtgtgtaca acatgatggg aaaaagagag 8940

aagaagctag gggaattcgg caaggcaaaa ggcagcagag ccatatggta catgtggctt 9000

ggagcacgct tcttagagtt tgaagcccta ggattcttaa atgaagatca ctggttctcc 9060

agagagaact ccctgagtgg agtggaagga gaagggctgc acaagctagg ttacattcta 9120

agagacgtga gcaagaaaga gggaggagca atgtatgccg atgacaccgc aggatgggat 9180

acaagaatca cactagaaga cctaaaaaat gaagaaatgg taacaaacca catggaagga 9240

gaacacaaga aactagccga ggccattttc aaactaacgt accaaaacaa ggtggtgcgt 9300

gtgcaaagac caacaccaag aggcacagta atggacatca tatcgagaag agaccaaaga 9360

ggtagtggac aagttggcac ctatggactc aatactttca ccaatatgga agcccaacta 9420

atcagacaga tggagggaga aggagtcttt aaaagcattc agcacctaac aatcacagaa 9480

gaaatcgctg tgcaaaactg gttagcaaga gtggggcgcg aaaggttatc aagaatggcc 9540

atcagtggag atgattgtgt tgtgaaacct ttagatgaca ggttcgcaag cgctttaaca 9600

gctctaaatg acatgggaaa gattaggaaa gacatacaac aatgggaacc ttcaagagga 9660

tggaatgatt ggacacaagt gcccttctgt tcacaccatt tccatgagtt aatcatgaaa 9720

gacggtcgcg tactcgttgt tccatgtaga aaccaagatg aactgattgg cagagcccga 9780

atctcccaag gagcagggtg gtctttgcgg gagacggcct gtttggggaa gtcttacgcc 9840

caaatgtgga gcttgatgta cttccacaga cgcgacctca ggctggcggc aaatgctatt 9900

tgctcggcag taccatcaca ttgggttcca acaagtcgaa caacctggtc catacatgct 9960

aaacatgaat ggatgacaac ggaagacatg ctgacagtct ggaacagggt gtggattcaa 10020

gaaaacccat ggatggaaga caaaactcca gtggaatcat gggaggaaat cccatacttg 10080

gggaaaagag aagaccaatg gtgcggctca ttgattgggt taacaagcag ggccacctgg 10140

gcaaagaaca tccaagcagc aataaatcaa gttagatccc ttataggcaa tgaagaatac 10200

acagattaca tgccatccat gaaaagattc agaagagaag aggaagaagc aggagttctg 10260

tggtagaaag caaaactaac atgaaacaag gctagaagtc aggtcggatt aagccatagt 10320

acggaaaaaa ctatgctacc tgtgagcccc gtccaaggac gttaaaagaa gtcaggccat 10380

cataaatgcc atagcttgag taaactatgc agcctgtagc tccacctgag aaggtgtaaa 10440

aaatccggga ggccacaaac catggaagct gtacgcatgg cgtagtggac tagcggttag 10500

aggagacccc tcccttacaa atcgcagcaa caatgggggc ccaaggcgag atgaagctgt 10560

agtctcgctg gaaggactag aggttagagg agacccccccc gaaacaaaaa acagcatatt 10620

gacgctggga aagaccagag atcctgctgt ctcctcagca tcattccagg cacagaacgc 10680

cagaaaatgg aatggtgctg ttgaatcaac aggttct 10717

<210> 20
<211> 3389
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 3, MVS

<400> 20

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Arg Met Ile Val
115 120 125

```
Gly Lys Asn Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ala Ser Gly
    130                 135                 140

Ile Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Asp
145                 150                 155                 160

Asp Thr Val Thr Tyr Lys Cys Pro His Ile Thr Glu Val Glu Pro Glu
                165                 170                 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Thr Tyr Gly
            180                 185                 190

Thr Cys Asn Gln Ala Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
        195                 200                 205

Leu Ala Pro His Val Gly Met Gly Leu Asp Thr Arg Thr Gln Thr Trp
    210                 215                 220

Met Ser Ala Glu Gly Ala Trp Arg Gln Val Glu Lys Val Glu Thr Trp
225                 230                 235                 240

Ala Leu Arg His Pro Gly Phe Thr Ile Leu Ala Leu Phe Leu Ala His
                245                 250                 255

Tyr Ile Gly Thr Ser Leu Thr Gln Lys Val Val Ile Phe Ile Leu Leu
            260                 265                 270

Met Leu Val Thr Pro Ser Met Thr Met Arg Cys Val Gly Val Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
    290                 295                 300

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Ile Glu Leu Gln Lys Thr Glu Ala Thr Gln Leu Ala Thr Leu
                325                 330                 335
```

```
Arg Lys Leu Cys Ile Glu Gly Lys Ile Thr Asn Ile Thr Thr Asp Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Ala Ile Leu Pro Glu Glu Gln Asp Gln
        355                 360                 365

Asn Tyr Val Cys Lys His Thr Tyr Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Val Thr Cys Ala Lys Phe Gln
385                 390                 395                 400

Cys Leu Glu Ser Ile Glu Gly Lys Val Val Gln His Glu Asn Leu Lys
                405                 410                 415

Tyr Thr Val Ile Ile Thr Val His Thr Gly Asp Gln His Gln Val Gly
            420                 425                 430

Asn Glu Thr Gln Gly Val Thr Ala Glu Ile Thr Pro Gln Ala Ser Thr
        435                 440                 445

Ala Glu Ala Ile Leu Pro Glu Tyr Gly Thr Leu Gly Leu Glu Cys Ser
    450                 455                 460

Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Ile Ser Leu Thr Met Lys
465                 470                 475                 480

Asn Lys Ala Trp Met Val His Arg Gln Trp Phe Phe Asp Leu Pro Leu
                485                 490                 495

Pro Trp Thr Ser Gly Ala Ser Ala Glu Thr Pro Thr Trp Asn Arg Lys
            500                 505                 510

Glu Leu Leu Val Thr Phe Lys Asn Ala His Ala Lys Lys Gln Glu Val
        515                 520                 525

Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu Thr Gly
    530                 535                 540
```

Ala Thr Glu Ile Gln Thr Ser Gly Gly Thr Ser Ile Phe Ala Gly His
545 550 555 560

Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Glu Leu Lys Gly Met Ser
565 570 575

Tyr Ala Met Cys Leu Ser Ser Phe Val Leu Lys Lys Glu Val Ser Glu
580 585 590

Thr Gln His Gly Thr Ile Leu Ile Lys Val Glu Tyr Lys Gly Glu Asp
595 600 605

Ala Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp Gly Gln Gly Lys Ala
610 615 620

Leu Asn Gly Arg Leu Ile Thr Ala Asn Pro Val Val Thr Lys Lys Glu
625 630 635 640

Glu Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser Asn Ile
645 650 655

Val Ile Gly Ile Gly Asp Lys Ala Leu Lys Ile Asn Trp Tyr Lys Lys
660 665 670

Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly Ala Arg
675 680 685

Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly
690 695 700

Gly Val Leu Asn Ser Leu Gly Lys Met Val His Gln Ile Phe Gly Ser
705 710 715 720

Ala Tyr Thr Ala Leu Phe Gly Gly Val Ser Trp Met Met Lys Ile Gly
725 730 735

Ile Gly Val Leu Leu Thr Trp Ile Gly Leu Asn Ser Lys Asn Thr Ser
740 745 750

Met Ser Phe Ser Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr Leu Gly
755 760 765

Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys Asn Lys
770 775 780

Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val His Thr
785 790 795 800

Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys Leu Ala
805 810 815

Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile Arg Ser
820 825 830

Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro Glu Leu
835 840 845

Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met Thr Gly
850 855 860

Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg Pro Gln
865 870 875 880

Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala Lys Met
885 890 895

Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly Pro Glu
900 905 910

Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu Glu Val
915 920 925

Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu Lys Leu
930 935 940

Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser Ala Ala
945 950 955 960

Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp Ile Glu
965 970 975

Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe Ile Glu
980 985 990

Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser Asn Gly
995 1000 1005

Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly Pro
1010 1015 1020

Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
1025 1030 1035

Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys
1040 1045 1050

Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly
1055 1060 1065

Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu
1070 1075 1080

Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly
1085 1090 1095

Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu
1100 1105 1110

Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly His Gly
1115 1120 1125

Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala Leu Phe
1130 1135 1140

Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His Ala Ile
1145 1150 1155

```
Leu Leu  Val Ala Val Ser Phe  Val Thr Leu Ile Thr  Gly Asn Met
    1160                1165                1170

Ser Phe  Arg Asp Leu Gly Arg  Val Met Val Met Val  Gly Ala Thr
    1175                1180                1185

Met Thr  Asp Asp Ile Gly Met  Gly Val Thr Tyr Leu  Ala Leu Leu
    1190                1195                1200

Ala Ala  Phe Lys Val Arg Pro  Thr Phe Ala Ala Gly  Leu Leu Leu
    1205                1210                1215

Arg Lys  Leu Thr Ser Lys Glu  Leu Met Met Thr Thr  Ile Gly Ile
    1220                1225                1230

Val Leu  Leu Ser Gln Ser Thr  Ile Pro Glu Thr Ile  Leu Glu Leu
    1235                1240                1245

Thr Asp  Ala Leu Ala Leu Gly  Met Met Val Leu Lys  Met Val Arg
    1250                1255                1260

Asn Met  Glu Lys Tyr Gln Leu  Ala Val Thr Ile Met  Ala Ile Leu
    1265                1270                1275

Cys Val  Pro Asn Ala Val Ile  Leu Gln Asn Ala Trp  Lys Val Ser
    1280                1285                1290

Cys Thr  Ile Leu Ala Val Val  Ser Val Ser Pro Leu  Phe Leu Thr
    1295                1300                1305

Ser Ser  Gln Gln Lys Thr Asp  Trp Ile Pro Leu Ala  Leu Thr Ile
    1310                1315                1320

Lys Gly  Leu Asn Pro Thr Ala  Ile Phe Leu Thr Thr  Leu Ser Arg
    1325                1330                1335

Thr Ser  Lys Lys Arg Ser Trp  Pro Leu Asn Glu Ala  Ile Met Ala
    1340                1345                1350
```

```
Val Gly  Met Val Ser Ile Leu  Ala Ser Ser Leu Leu  Lys Asn Asp
    1355                 1360                 1365

Ile Pro  Met Thr Gly Pro Leu  Val Ala Gly Gly Leu  Leu Thr Val
    1370                 1375                 1380

Cys Tyr  Val Leu Thr Gly Arg  Ser Ala Asp Leu Glu  Leu Glu Arg
    1385                 1390                 1395

Ala Ala  Asp Val Lys Trp Glu  Asp Gln Ala Glu Ile  Ser Gly Ser
    1400                 1405                 1410

Ser Pro  Ile Leu Ser Ile Thr  Ile Ser Glu Asp Gly  Ser Met Ser
    1415                 1420                 1425

Ile Lys  Asn Glu Glu Glu Glu  Gln Thr Leu Thr Ile  Leu Ile Arg
    1430                 1435                 1440

Thr Gly  Leu Leu Val Ile Ser  Gly Leu Phe Pro Val  Ser Ile Pro
    1445                 1450                 1455

Ile Thr  Ala Ala Ala Trp Tyr  Leu Trp Glu Val Lys  Lys Gln Arg
    1460                 1465                 1470

Ala Gly  Val Leu Trp Asp Val  Pro Ser Pro Pro Pro  Met Gly Lys
    1475                 1480                 1485

Ala Glu  Leu Glu Asp Gly Ala  Tyr Arg Ile Lys Gln  Lys Gly Ile
    1490                 1495                 1500

Leu Gly  Tyr Ser Gln Ile Gly  Ala Gly Val Tyr Lys  Glu Gly Thr
    1505                 1510                 1515

Phe His  Thr Met Trp His Val  Thr Arg Gly Ala Val  Leu Met His
    1520                 1525                 1530

Lys Gly  Lys Arg Ile Glu Pro  Ser Trp Ala Asp Val  Lys Lys Asp
    1535                 1540                 1545
```

```
Leu Ile  Ser Tyr Gly Gly Gly  Trp Lys Leu Glu Gly  Glu Trp Lys
    1550                1555                1560

Glu Gly  Glu Glu Val Gln Val  Leu Ala Leu Glu Pro  Gly Lys Asn
    1565                1570                1575

Pro Arg  Ala Val Gln Thr Lys  Pro Gly Leu Phe Lys  Thr Asn Ala
    1580                1585                1590

Gly Thr  Ile Gly Ala Val Ser  Leu Asp Phe Ser Pro  Gly Thr Ser
    1595                1600                1605

Gly Ser  Pro Ile Ile Asp Lys  Lys Gly Lys Val Val  Gly Leu Tyr
    1610                1615                1620

Gly Asn  Gly Val Val Thr Arg  Ser Gly Ala Tyr Val  Ser Ala Ile
    1625                1630                1635

Ala Gln  Thr Glu Lys Ser Ile  Glu Asp Asn Pro Glu  Ile Glu Asp
    1640                1645                1650

Asp Ile  Phe Arg Lys Arg Arg  Leu Thr Ile Met Asp  Leu His Pro
    1655                1660                1665

Gly Ala  Gly Lys Thr Lys Arg  Tyr Leu Pro Ala Ile  Val Arg Glu
    1670                1675                1680

Ala Ile  Lys Arg Gly Leu Arg  Thr Leu Ile Leu Ala  Pro Thr Arg
    1685                1690                1695

Val Val  Ala Ala Glu Met Glu  Glu Ala Leu Arg Gly  Leu Pro Ile
    1700                1705                1710

Arg Tyr  Gln Thr Pro Ala Ile  Arg Ala Val His Thr  Gly Arg Glu
    1715                1720                1725

Ile Val  Asp Leu Met Cys His  Ala Thr Phe Thr Met  Arg Leu Leu
    1730                1735                1740
```

Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
1745 1750 1755

Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile
1760 1765 1770

Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr
1775 1780 1785

Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala
1790 1795 1800

Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn
1805 1810 1815

Ser Gly His Glu Trp Val Thr Asp Phe Lys Gly Lys Thr Val Trp
1820 1825 1830

Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala Cys Leu
1835 1840 1845

Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg Lys Thr Phe
1850 1855 1860

Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp Phe Val
1865 1870 1875

Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys Ala Glu
1880 1885 1890

Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile Leu Thr
1895 1900 1905

Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro Val Thr
1910 1915 1920

His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg Asn Pro
1925 1930 1935

Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro Leu Glu
1940 1945 1950

Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met Leu Leu
1955 1960 1965

Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe Glu
1970 1975 1980

Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu
1985 1990 1995

Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly
2000 2005 2010

Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile
2015 2020 2025

Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn
2030 2035 2040

Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu
2045 2050 2055

Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile
2060 2065 2070

Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala
2075 2080 2085

Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met Gly Arg
2090 2095 2100

Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu Asp Asn
2105 2110 2115

Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala Tyr Asn
2120 2125 2130

His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu Leu Leu
2135 2140 2145

Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu Phe Leu
2150 2155 2160

Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met Cys Cys
2165 2170 2175

Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile Gln Pro
2180 2185 2190

His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu Ile Val
2195 2200 2205

Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp Asn
2210 2215 2220

Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val Ala Ala
2225 2230 2235

Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp
2240 2245 2250

Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile
2255 2260 2265

Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala
2270 2275 2280

Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu
2285 2290 2295

Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala
2300 2305 2310

Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser Lys Met
2315 2320 2325

Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser Gln Val
2330 2335 2340

Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val Ala His
2345 2350 2355

Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr Arg Glu
2360 2365 2370

Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro Thr Val
2375 2380 2385

Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr Asp Pro
2390 2395 2400

Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val Leu Cys
2405 2410 2415

Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu Cys Glu
2420 2425 2430

Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly
2435 2440 2445

Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala
2450 2455 2460

Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe
2465 2470 2475

Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn
2480 2485 2490

Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala
2495 2500 2505

Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln
2510 2515 2520

Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu
2525 2530 2535

Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp
2540 2545 2550

Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val Val Asp
2555 2560 2565

Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly Gly Leu
2570 2575 2580

Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly Pro Gly
2585 2590 2595

His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn Leu Val
2600 2605 2610

Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu Lys
2615 2620 2625

Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro Asn Pro
2630 2635 2640

Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu Val Glu
2645 2650 2655

Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val Leu Asn
2660 2665 2670

Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu Gln Arg
2675 2680 2685

Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn Ser
2690 2695 2700

Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val
2705 2710 2715

Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr
2720 2725 2730

Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly
2735 2740 2745

Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu
2750 2755 2760

Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu His Glu
2765 2770 2775

Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr Trp Ala
2780 2785 2790

Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala Ser Ser
2795 2800 2805

Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp Asp Val
2810 2815 2820

Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr Pro Phe
2825 2830 2835

Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg Thr Gln
2840 2845 2850

Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr Ala Glu
2855 2860 2865

Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg Met Cys
2870 2875 2880

Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala Ala Leu
2885 2890 2895

Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala Arg Glu
2900 2905 2910

Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys Glu Arg
2915 2920 2925

Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn Met
2930 2935 2940

Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
2945 2950 2955

Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu
2960 2965 2970

Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser
2975 2980 2985

Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys
2990 2995 3000

Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly Gly Ala
3005 3010 3015

Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Leu
3020 3025 3030

Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met Glu Gly
3035 3040 3045

Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr Tyr Gln
3050 3055 3060

Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly Thr Val
3065 3070 3075

Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly Gln Val
3080 3085 3090

Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala Gln Leu
3095 3100 3105

Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile Gln His
3110 3115 3120

Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu Ala Arg
3125 3130 3135

Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly Asp Asp
3140 3145 3150

Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala Leu Thr
3155 3160 3165

Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln Trp
3170 3175 3180

Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys
3185 3190 3195

Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu
3200 3205 3210

Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg
3215 3220 3225

Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu
3230 3235 3240

Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg
3245 3250 3255

Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro
3260 3265 3270

Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile His Ala
3275 3280 3285

Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val Trp Asn
3290 3295 3300

Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys Thr Pro
3305 3310 3315

Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu Asp
3320 3325 3330

Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp
3335 3340 3345

Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile
3350 3355 3360

Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe
3365 3370 3375

Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385

<210> 21
<211> 10717
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 3, WVS

<400> 21
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgatg gagagccgcg catgattgtg 480

gggaagaatg aaagaggaaa atccctactt ttcaagacag cctctggaat caacatgtgc 540

acactcatag ctatggatct gggagagatg tgtgatgaca cggtcactta caaatgcccc 600

cacattaccg aagtggagcc tgaagacatt gactgctggt gcaaccttac atcgacatgg 660

gtgacttatg gaacatgcaa tcaagctgga gagcatagac gcgataagag atcagtggcg 720

ttagctcccc atgttggcat gggactggac acacgcactc aaacctggat gtcggctgaa 780

ggagcttgga gacaagtcga gaaggtagag acatgggccc ttaggcaccc agggtttacc 840

atactagccc tatttcttgc ccattacata ggcacttcct tgacccagaa agtggttatt 900

tttatactat taatgctggt taccccatcc atgacaatga gatgtgtagg agtaggaaac 960

agagattttg tggaaggcct atcgggagct acgtgggttg acgtggtgct cgagcacggt 1020

gggtgtgtga ctaccatggc taagaacaag cccacgctgg acatagagct tcagaagacc 1080

gaggccaccc aactggcgac cctaaggaag ctatgcattg agggaaaaat taccaacata 1140

acaaccgact caagatgtcc cacccaaggg gaagcgattt tacctgagga gcaggaccag 1200

aactacgtgt gtaagcatac atacgtggac agaggctggg gaaacggttg tggtttgttt 1260

ggcaagggaa gcttggtgac atgcgcgaaa tttcaatgtt tagaatcaat agagggaaaa 1320

gtggtgcaac atgagaacct caaatacacc gtcatcatca cagtgcacac aggagaccaa 1380

caccaggtgg gaaatgaaac gcagggagtc acggctgaga taacacccca ggcatcaacc 1440

gctgaagcca ttttacctga atatggaacc ctcgggctag aatgctcacc acggacaggt 1500

ttggatttca atgaaatgat ctcattgaca atgaagaaca aagcatggat ggtacataga 1560

caatggttct ttgacttacc cctaccatgg acatcaggag cttcagcaga aacaccaact 1620

tggaacagga aagagcttct tgtgacattt aaaaatgcac atgcaaaaaa gcaagaagta 1680

gttgttcttg gatcacaaga gggagcaatg catacagcac tgacaggagc tacagagatc 1740

caaacctcag gaggcacaag tatctttgcg gggcacttaa aatgtagact caagatggac 1800

aaattggaac tcaaggggat gagctatgca atgtgcttga gtagctttgt gttgaagaaa 1860

gaagtctccg aaacgcagca tgggacaata ctcattaagg ttgagtacaa aggggaagat 1920

gcaccctgca agattccttt ctccacggag gatggacaag gaaaagctct caatggcaga 1980

ctgatcacag ccaatccagt ggtgaccaag aaggaggagc ctgtcaacat tgaggctgaa 2040

cctccttttg gagaaagtaa catagtaatt ggaattggag acaaagccct gaaaatcaac 2100

tggtacaaga agggaagctc gattgggaag atgttcgagg ccactgccag aggtgcaagg 2160

cgcatggcca tcttgggaga cacagcctgg gactttggat cagtgggtgg tgttttgaat 2220

tcattaggga aaatggtcca ccaaatattt gggagtgctt acacagccct atttggtgga 2280

gtctcctgga tgatgaaaat tggaataggt gtcctcttaa cctggatagg gttgaactca 2340

aaaaatactt ctatgtcatt ttcatgcatc gcggccggca ttgtgacact gtatttggga 2400

gtcatggtgc aggccgatag tggttgcgtt gtgagctgga aaaacaaaga actgaaatgt 2460

ggcagtggga ttttcatcac agacaacgtg cacacatgga cagaacaata caagttccaa 2520

ccagaatccc cttcaaaact agcttcagct atccagaaag cccatgaaga ggacatttgt 2580

ggaatccgct cagtaacaag actggagaat ctgatgtgga aacaaataac accagaattg 2640

aatcacattc tatcagaaaa tgaggtgaag ttaactatta tgacaggaga catcaaagga 2700

atcatgcagg caggaaaacg atctctgcgg cctcagccca ctgagctgaa gtattcatgg 2760

aaaacatggg gcaaagcaaa aatgctctct acagagtctc ataaccagac ctttctcatt 2820

gatggccccg aaacagcaga atgccccaac acaaatagag cttggaattc gttggaagtt 2880

gaagactatg gctttggagt attcaccacc aatatatggc taaaattgaa agaaaaacag 2940

gatgtattct gcgactcaaa actcatgtca gcggccataa aagacaacag agccgtccat 3000

gccgatatgg gttattggat agaaagtgca ctcaatgaca catggaagat agagaaagcc 3060

tctttcattg aagttaaaaa ctgccactgg ccaaaatcac acaccctctg gagcaatgga 3120

gtgctagaaa gtgagatgat aattccaaag aatctcgctg gaccagtgtc tcaacacaac 3180

tatagaccag gctaccatac acaaataaca ggaccatggc atctaggtaa gcttgagatg 3240

gactttgatt tctgtgatgg aacaacagtg gtagtgactg aggactgcgg aaatagagga 3300

ccctctttga gaacaaccac tgcctctgga aaactcataa cagaatggtg ctgccgatct 3360

tgcacattac caccgctaag atacagaggt gaggatgggt gctggtacgg gatggaaatc 3420

agaccattga aggagaaaga agagaatttg gtcaactcct tggtcacagc tggacatggg 3480

caggtcgaca acttttcact aggagtcttg ggaatggcat tgttcctgga ggaaatgctt 3540

aggacccgag taggaacgaa acatgcaata ctactagttg cagtttcttt tgtgacattg 3600

atcacaggga acatgtcctt tagagacctg ggaagagtga tggttatggt aggcgccact 3660

atgacggatg acataggtat gggcgtgact tatcttgccc tactagcagc cttcaaagtc 3720

agaccaactt ttgcagctgg actactcttg agaaagctga cctccaagga attgatgatg 3780

actactatag gaattgtact cctctcccag agcaccatac cagagaccat tcttgagttg 3840

actgatgcgt tagccttagg catgatggtc ctcaaaatgg tgagaaatat ggaaaagtat 3900

caattggcag tgactatcat ggctatcttg tgcgtcccaa acgcagtgat attacaaaac 3960

gcatggaaag tgagttgcac aatattggca gtggtgtccg tttccccact gttcttaaca 4020

tcctcacagc aaaaaacaga ttggatacca ttagcattga cgatcaaagg tctcaatcca 4080

acagctattt ttctaacaac cctctcaaga accagcaaga aaaggagctg gccattaaat 4140

gaggctatca tggcagtcgg gatggtgagc attttagcca gttctctcct aaaaaatgat 4200

attcccatga caggaccatt agtggctgga gggctcctca ctgtgtgcta cgtgctcact 4260

ggacgatcgg ccgatttgga actggagaga gcagccgatg tcaaatggga agaccaggca 4320

gagatatcag gaagcagtcc aatcctgtca ataacaatat cagaagatgg tagcatgtcg 4380

ataaaaaatg aagaggaaga acaaacactg accatactca ttagaacagg attgctggtg 4440

atctcaggac tttttcctgt atcaatacca atcacggcag cagcatggta cctgtgggaa 4500

gtgaagaaac aacgggccgg agtattgtgg gatgttcctt cacccccacc catgggaaag 4560

gctgaactgg aagatggagc ctatagaatt aagcaaaaag ggattcttgg atattcccag 4620

atcggagccg gagtttacaa agaaggaaca ttccatacaa tgtggcatgt cacacgtggc 4680

gctgttctaa tgcataaagg aaagaggatt gaaccatcat gggcggacgt caagaaagac 4740

ctaatatcat atggaggagg ctggaagtta gaaggagaat ggaaggaagg agaagaagtc 4800

caggtattgg cactggagcc tggaaaaaat ccaagagccg tccaaacgaa acctggtctt 4860

ttcaaaacca acgccggaac aataggtgct gtatctctgg acttttctcc tggaacgtca 4920

ggatctccaa ttatcgacaa aaaaggaaaa gttgtgggtc tttatggtaa tggtgttgtt 4980

acaaggagtg gagcatatgt gagtgctata gcccagactg aaaaaagcat tgaagacaac 5040

ccagagatcg aagatgacat tttccgaaag agaagactga ccatcatgga cctccaccca 5100

ggagcgggaa agacgaagag ataccttccg gccatagtca gagaagctat aaaacggggt 5160

ttgagaacat taatcttggc ccccactaga gttgtggcag ctgaaatgga ggaagccctt 5220

agaggacttc caataagata ccagacccca gccatcagag ctgtgcacac cgggcgggag 5280

attgtggacc taatgtgtca tgccacattt accatgaggc tgctatcacc agttagagtg 5340

ccaaactaca acctgattat catggacgaa gcccatttca cagacccagc aagtatagca 5400

gctagaggat acatctcaac tcgagtggag atgggtgagg cagctgggat ttttatgaca 5460

gccactcccc cgggaagcag agacccattt cctcagagca atgcaccaat catagatgaa 5520

gaaagagaaa tccctgaacg ctcgtggaat tccggacatg aatgggtcac ggatttaaa 5580

gggaagactg tttggttcgt tccaagtata aaagcaggaa atgatatagc agcttgcctg 5640

aggaaaaatg gaaagaaagt gatacaactc agtaggaaga cctttgattc tgagtatgtc 5700

aagactagaa ccaatgattg ggacttcgtg gttacaactg acatttcaga aatgggtgcc 5760

aatttcaagg ctgagagggt tatagacccc agacgctgca tgaaaccagt catactaaca 5820

gatggtgaag agcgggtgat tctggcagga cctatgccag tgacccactc tagtgcagca 5880

caaagaagag ggagaatagg aagaaatcca aaaaatgaga atgaccagta catatacatg 5940

ggggaacctc tggaaaatga tgaagactgt gcacactgga aagaagctaa aatgctccta 6000

gataacatca acacgccaga aggaatcatt cctagcatgt tcgaaccaga gcgtgaaaag 6060

gtggatgcca ttgatggcga ataccgcttg agaggagaag caaggaaaac ctttgtagac 6120

ttaatgagaa gaggagacct accagtctgg ttggcctaca gagtggcagc tgaaggcatc 6180

aactacgcag acagaaggtg gtgttttgat ggagtcaaga acaaccaaat cctagaagaa 6240

aacgtggaag ttgaaatctg gacaaaagaa ggggaaagga agaaattgaa acccagatgg 6300

ttggatgcta ggatctattc tgacccactg gcgctaaaag aatttaagga atttgcagcc 6360

ggaagaaagt ctctgaccct gaacctaatc acagaaatgg gtaggctccc aaccttcatg 6420

actcagaagg caagaaacgc actggacaac ttagcagtgc tgcacacggc tgaggcaggt 6480

ggaagggcgt acaaccatgc tctcagtgaa ctgccggaga ccctggagac attgctttta 6540

ctgacacttc tggctacagt cacgggaggg atctttttat tcttgatgag cgcaaggggc 6600

atagggaaga tgaccctggg aatgtgctgc ataatcacgg ctagcatcct cctatggtac 6660

gcacaaatac agccacactg gatagcagct tcaataatac tggagttttt tctcatagtt 6720

ttgcttattc cagaacctga aaaacagaga acaccccaag acaaccaact gacctacgtt 6780

gtcatagcca tcctcacagt ggtggccgca accatggcaa acgagatggg tttcctagaa 6840

aaaacgaaga aagatctcgg attgggaagc attgcaaccc agcaacccga gagcaacatc 6900

ctggacatag atctacgtcc tgcatcagca tggacgctgt atgccgtggc cacaacattt 6960

gttacaccaa tgttgagaca tagcattgaa aattcctcag tgaatgtgtc cctaacagct 7020

atagccaacc aagccacagt gttaatgggt ctcgggaaag gatggccatt gtcaaagatg 7080

gacatcggag ttccccttct cgccattgga tgctactcac aagtcaaccc cataactctc 7140

acagcagctc ttttcttatt ggtagcacat tatgccatca tagggccagg actccaagca 7200

aaagcaacca gagaagctca gaaaagagca gcggcgggca tcatgaaaaa cccaactgtc 7260

gatggaataa cagtgattga cctagatcca ataccttatg atccaaagtt tgaaaagcag 7320

ttgggacaag taatgctcct agtcctctgc gtgactcaag tattgatgat gaggactaca 7380

tgggctctgt gtgaggcttt aaccttagct accgggccca tctccacatt gtgggaagga 7440

aatccaggga ggttttggaa cactaccatt gcggtgtcaa tggctaacat ttttagaggg 7500

agttacttgg ccggagctgg acttctcttt tctattatga agaacacaac caacacaaga 7560

aggggaactg gcaacatagg agagacgctt ggagagaaat ggaaaagccg attgaacgcg 7620

ttgggaaaaa gtgaattcca gatctacaag aaaagtggaa tccaggaagt ggatagaacc 7680

ttagcaaaag aaggcattaa aagaggagaa acggaccatc acgctgtgtc gcgaggctca 7740

gcaaaactga gatggttcgt tgagagaaac atggtcacac cagaagggaa agtagtggac 7800

ctcggttgtg gcagaggagg ctggtcatac tattgtggag gactaaagaa tgtaagagaa 7860

gtcaaaggcc taacaaaagg aggaccagga cacgaagaac ccatccccat gtcaacatat 7920

gggtggaatc tagtgcgtct tcaaagtgga gttgacgttt tcttcatccc gccagaaaag 7980

tgtgacacat tattgtgtga cataggggag tcatcaccaa atcccacagt ggaagcagga 8040

cgaacactca gagtccttaa cttagtagaa aattggttga acaacaacac tcaattttgc 8100

ataaaggttc tcaacccata tatgccctca gtcatagaaa aaatggaagc actacaaagg 8160

aaatatggag gagccttagt gaggaatcca ctctcacgaa actccacaca tgagatgtac 8220

tgggtatcca atgcttccgg gaacatagtg tcatcagtga acatgatttc aaggatgttg 8280

atcaacagat ttacaatgag atacaagaaa gccacttacg agccggatgt tgacctcgga 8340

agcggaaccc gtaacatcgg gattgaaagt gagataccaa acctagatat aattgggaaa 8400

agaatagaaa aaataaagca agagcatgaa acatcatggc actatgacca agaccaccca 8460

tacaaaacgt gggcatacca tggtagctat gaaacaaaac agactggatc agcatcatcc 8520

atggtcaacg gagtggtcag gctgctgaca aaaccttggg acgtcgtccc catggtgaca 8580

cagatggcaa tgacagacac gactccattt ggacaacagc gcgtttttaa agagaaagtg 8640

gacacgagaa cccaagaacc gaaagaaggc acgaagaaac taatgaaaat aacagcagag 8700

tggctttgga aagaattagg gaagaaaaag acacccagga tgtgcaccag agaagaattc 8760

acaagaaagg tgagaagcaa tgcagccttg ggggccatat tcactgatga gaacaagtgg 8820

aagtcggcac gtgaggctgt tgaagatagt aggttttggg agctggttga caaggaaagg 8880

aatctccatc ttgaaggaaa gtgtgaaaca tgtgtgtaca acatgatggg aaaaagagag 8940

aagaagctag gggaattcgg caaggcaaaa ggcagcagag ccatatggta catgtggctt 9000

ggagcacgct tcttagagtt tgaagcccta ggattcttaa atgaagatca ctggttctcc 9060

agagagaact ccctgagtgg agtggaagga gaagggctgc acaagctagg ttacattcta 9120

agagacgtga gcaagaaaga gggaggagca atgtatgccg atgacaccgc aggatgggat 9180

acaagaatca cactagaaga cctaaaaaat gaagaaatgg taacaaacca catggaagga 9240

gaacacaaga aactagccga ggccattttc aaactaacgt accaaaacaa ggtggtgcgt 9300

gtgcaaagac caacaccaag aggcacagta atggacatca tatcgagaag agaccaaaga 9360

ggtagtggac aagttggcac ctatggactc aatactttca ccaatatgga agcccaacta 9420

atcagacaga tggagggaga aggagtcttt aaaagcattc agcacctaac aatcacagaa 9480

gaaatcgctg tgcaaaactg gttagcaaga gtggggcgcg aaaggttatc aagaatggcc 9540

atcagtggag atgattgtgt tgtgaaacct ttagatgaca ggttcgcaag cgctttaaca 9600

gctctaaatg acatgggaaa gattaggaaa gacatacaac aatgggaacc ttcaagagga 9660

tggaatgatt ggacacaagt gcccttctgt tcacaccatt tccatgagtt aatcatgaaa 9720

gacggtcgcg tactcgttgt tccatgtaga aaccaagatg aactgattgg cagagcccga 9780

atctcccaag gagcagggtg gtctttgcgg gagacggcct gtttggggaa gtcttacgcc 9840

caaatgtgga gcttgatgta cttccacaga cgcgacctca ggctggcggc aaatgctatt 9900

tgctcggcag taccatcaca ttgggttcca acaagtcgaa caacctggtc catacatgct 9960

aaacatgaat ggatgacaac ggaagacatg ctgacagtct ggaacagggt gtggattcaa 10020

gaaaacccat ggatggaaga caaaactcca gtggaatcat gggaggaaat cccatacttg 10080

gggaaaagag aagaccaatg gtgcggctca ttgattgggt taacaagcag ggccacctgg 10140

gcaaagaaca tccaagcagc aataaatcaa gttagatccc ttataggcaa tgaagaatac 10200

acagattaca tgccatccat gaaaagattc agaagagaag aggaagaagc aggagttctg 10260

tggtagaaag caaaactaac atgaaacaag gctagaagtc aggtcggatt aagccatagt 10320

acggaaaaaa ctatgctacc tgtgagcccc gtccaaggac gttaaaagaa gtcaggccat 10380

cataaatgcc atagcttgag taaactatgc agcctgtagc tccacctgag aaggtgtaaa 10440

aaatccggga ggccacaaac catggaagct gtacgcatgg cgtagtggac tagcggttag 10500

aggagacccc tcccttacaa atcgcagcaa caatgggggc ccaaggcgag atgaagctgt 10560

agtctcgctg gaaggactag aggttagagg agacccccc gaaacaaaaa acagcatatt 10620

gacgctggga aagaccagag atcctgctgt ctcctcagca tcattccagg cacagaacgc 10680

cagaaaatgg aatggtgctg ttgaatcaac aggttct 10717

<210> 22
<211> 3389
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 3, WVS

<400> 22

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Arg Met Ile Val
115 120 125

Gly Lys Asn Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ala Ser Gly
130 135 140

Ile Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Asp
145 150 155 160

Asp Thr Val Thr Tyr Lys Cys Pro His Ile Thr Glu Val Glu Pro Glu
165 170 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Thr Tyr Gly
180 185 190

Thr Cys Asn Gln Ala Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
195 200 205

Leu Ala Pro His Val Gly Met Gly Leu Asp Thr Arg Thr Gln Thr Trp
210 215 220

Met Ser Ala Glu Gly Ala Trp Arg Gln Val Glu Lys Val Glu Thr Trp
225 230 235 240

Ala Leu Arg His Pro Gly Phe Thr Ile Leu Ala Leu Phe Leu Ala His
245 250 255

Tyr Ile Gly Thr Ser Leu Thr Gln Lys Val Val Ile Phe Ile Leu Leu
260 265 270

```
Met Leu Val Thr Pro Ser Met Thr Met Arg Cys Val Gly Val Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
    290                 295                 300

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Ile Glu Leu Gln Lys Thr Glu Ala Thr Gln Leu Ala Thr Leu
                325                 330                 335

Arg Lys Leu Cys Ile Glu Gly Lys Ile Thr Asn Ile Thr Thr Asp Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Ala Ile Leu Pro Glu Glu Gln Asp Gln
        355                 360                 365

Asn Tyr Val Cys Lys His Thr Tyr Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Val Thr Cys Ala Lys Phe Gln
385                 390                 395                 400

Cys Leu Glu Ser Ile Glu Gly Lys Val Val Gln His Glu Asn Leu Lys
                405                 410                 415

Tyr Thr Val Ile Ile Thr Val His Thr Gly Asp Gln His Gln Val Gly
            420                 425                 430

Asn Glu Thr Gln Gly Val Thr Ala Glu Ile Thr Pro Gln Ala Ser Thr
        435                 440                 445

Ala Glu Ala Ile Leu Pro Glu Tyr Gly Thr Leu Gly Leu Glu Cys Ser
    450                 455                 460

Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Ile Ser Leu Thr Met Lys
465                 470                 475                 480
```

Asn Lys Ala Trp Met Val His Arg Gln Trp Phe Phe Asp Leu Pro Leu
485 490 495

Pro Trp Thr Ser Gly Ala Ser Ala Glu Thr Pro Thr Trp Asn Arg Lys
500 505 510

Glu Leu Leu Val Thr Phe Lys Asn Ala His Ala Lys Lys Gln Glu Val
515 520 525

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

Ala Thr Glu Ile Gln Thr Ser Gly Gly Thr Ser Ile Phe Ala Gly His
545 550 555 560

Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Glu Leu Lys Gly Met Ser
565 570 575

Tyr Ala Met Cys Leu Ser Ser Phe Val Leu Lys Lys Glu Val Ser Glu
580 585 590

Thr Gln His Gly Thr Ile Leu Ile Lys Val Glu Tyr Lys Gly Glu Asp
595 600 605

Ala Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp Gly Gln Gly Lys Ala
610 615 620

Leu Asn Gly Arg Leu Ile Thr Ala Asn Pro Val Val Thr Lys Lys Glu
625 630 635 640

Glu Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser Asn Ile
645 650 655

Val Ile Gly Ile Gly Asp Lys Ala Leu Lys Ile Asn Trp Tyr Lys Lys
660 665 670

Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly Ala Arg
675 680 685

Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly
690 695 700

Gly Val Leu Asn Ser Leu Gly Lys Met Val His Gln Ile Phe Gly Ser
705 710 715 720

Ala Tyr Thr Ala Leu Phe Gly Gly Val Ser Trp Met Met Lys Ile Gly
725 730 735

Ile Gly Val Leu Leu Thr Trp Ile Gly Leu Asn Ser Lys Asn Thr Ser
740 745 750

Met Ser Phe Ser Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr Leu Gly
755 760 765

Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys Asn Lys
770 775 780

Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val His Thr
785 790 795 800

Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys Leu Ala
805 810 815

Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile Arg Ser
820 825 830

Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro Glu Leu
835 840 845

Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met Thr Gly
850 855 860

Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg Pro Gln
865 870 875 880

Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala Lys Met
885 890 895

Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly Pro Glu
900 905 910

Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu Glu Val
915 920 925

Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu Lys Leu
930 935 940

Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser Ala Ala
945 950 955 960

Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp Ile Glu
965 970 975

Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe Ile Glu
980 985 990

Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser Asn Gly
995 1000 1005

Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly Pro
1010 1015 1020

Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
1025 1030 1035

Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys
1040 1045 1050

Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly
1055 1060 1065

Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu
1070 1075 1080

Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly
1085 1090 1095

```
Glu Asp  Gly Cys Trp Tyr Gly  Met Glu Ile Arg Pro  Leu Lys Glu
    1100                1105                1110

Lys Glu  Glu Asn Leu Val Asn  Ser Leu Val Thr Ala  Gly His Gly
    1115                1120                1125

Gln Val  Asp Asn Phe Ser Leu  Gly Val Leu Gly Met  Ala Leu Phe
    1130                1135                1140

Leu Glu  Glu Met Leu Arg Thr  Arg Val Gly Thr Lys  His Ala Ile
    1145                1150                1155

Leu Leu  Val Ala Val Ser Phe  Val Thr Leu Ile Thr  Gly Asn Met
    1160                1165                1170

Ser Phe  Arg Asp Leu Gly Arg  Val Met Val Met Val  Gly Ala Thr
    1175                1180                1185

Met Thr  Asp Asp Ile Gly Met  Gly Val Thr Tyr Leu  Ala Leu Leu
    1190                1195                1200

Ala Ala  Phe Lys Val Arg Pro  Thr Phe Ala Ala Gly  Leu Leu Leu
    1205                1210                1215

Arg Lys  Leu Thr Ser Lys Glu  Leu Met Met Thr Thr  Ile Gly Ile
    1220                1225                1230

Val Leu  Leu Ser Gln Ser Thr  Ile Pro Glu Thr Ile  Leu Glu Leu
    1235                1240                1245

Thr Asp  Ala Leu Ala Leu Gly  Met Met Val Leu Lys  Met Val Arg
    1250                1255                1260

Asn Met  Glu Lys Tyr Gln Leu  Ala Val Thr Ile Met  Ala Ile Leu
    1265                1270                1275

Cys Val  Pro Asn Ala Val Ile  Leu Gln Asn Ala Trp  Lys Val Ser
    1280                1285                1290
```

Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe Leu Thr
1295 1300 1305

Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu Thr Ile
1310 1315 1320

Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu Ser Arg
1325 1330 1335

Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile Met Ala
1340 1345 1350

Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys Asn Asp
1355 1360 1365

Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu Thr Val
1370 1375 1380

Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu Glu Arg
1385 1390 1395

Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser Gly Ser
1400 1405 1410

Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser Met Ser
1415 1420 1425

Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu Ile Arg
1430 1435 1440

Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser Ile Pro
1445 1450 1455

Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys Gln Arg
1460 1465 1470

Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met Gly Lys
1475 1480 1485

```
Ala Glu  Leu Glu Asp Gly Ala  Tyr Arg Ile Lys Gln  Lys Gly Ile
    1490                 1495                 1500

Leu Gly  Tyr Ser Gln Ile Gly  Ala Gly Val Tyr Lys  Glu Gly Thr
    1505                 1510                 1515

Phe His  Thr Met Trp His Val  Thr Arg Gly Ala Val  Leu Met His
    1520                 1525                 1530

Lys Gly  Lys Arg Ile Glu Pro  Ser Trp Ala Asp Val  Lys Lys Asp
    1535                 1540                 1545

Leu Ile  Ser Tyr Gly Gly Gly  Trp Lys Leu Glu Gly  Glu Trp Lys
    1550                 1555                 1560

Glu Gly  Glu Glu Val Gln Val  Leu Ala Leu Glu Pro  Gly Lys Asn
    1565                 1570                 1575

Pro Arg  Ala Val Gln Thr Lys  Pro Gly Leu Phe Lys  Thr Asn Ala
    1580                 1585                 1590

Gly Thr  Ile Gly Ala Val Ser  Leu Asp Phe Ser Pro  Gly Thr Ser
    1595                 1600                 1605

Gly Ser  Pro Ile Ile Asp Lys  Lys Gly Lys Val Val  Gly Leu Tyr
    1610                 1615                 1620

Gly Asn  Gly Val Val Thr Arg  Ser Gly Ala Tyr Val  Ser Ala Ile
    1625                 1630                 1635

Ala Gln  Thr Glu Lys Ser Ile  Glu Asp Asn Pro Glu  Ile Glu Asp
    1640                 1645                 1650

Asp Ile  Phe Arg Lys Arg Arg  Leu Thr Ile Met Asp  Leu His Pro
    1655                 1660                 1665

Gly Ala  Gly Lys Thr Lys Arg  Tyr Leu Pro Ala Ile  Val Arg Glu
    1670                 1675                 1680
```

Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro Thr Arg
1685 1690 1695

Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu Pro Ile
1700 1705 1710

Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly Arg Glu
1715 1720 1725

Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu Leu
1730 1735 1740

Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
1745 1750 1755

Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile
1760 1765 1770

Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr
1775 1780 1785

Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala
1790 1795 1800

Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn
1805 1810 1815

Ser Gly His Glu Trp Val Thr Asp Phe Lys Gly Lys Thr Val Trp
1820 1825 1830

Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala Cys Leu
1835 1840 1845

Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg Lys Thr Phe
1850 1855 1860

Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp Phe Val
1865 1870 1875

```
Val Thr  Thr Asp Ile Ser Glu  Met Gly Ala Asn Phe  Lys Ala Glu
    1880                 1885                 1890

Arg Val  Ile Asp Pro Arg Arg  Cys Met Lys Pro Val  Ile Leu Thr
    1895                 1900                 1905

Asp Gly  Glu Glu Arg Val Ile  Leu Ala Gly Pro Met  Pro Val Thr
    1910                 1915                 1920

His Ser  Ser Ala Ala Gln Arg  Arg Gly Arg Ile Gly  Arg Asn Pro
    1925                 1930                 1935

Lys Asn  Glu Asn Asp Gln Tyr  Ile Tyr Met Gly Glu  Pro Leu Glu
    1940                 1945                 1950

Asn Asp  Glu Asp Cys Ala His  Trp Lys Glu Ala Lys  Met Leu Leu
    1955                 1960                 1965

Asp Asn  Ile Asn Thr Pro Glu  Gly Ile Ile Pro Ser  Met Phe Glu
    1970                 1975                 1980

Pro Glu  Arg Glu Lys Val Asp  Ala Ile Asp Gly Glu  Tyr Arg Leu
    1985                 1990                 1995

Arg Gly  Glu Ala Arg Lys Thr  Phe Val Asp Leu Met  Arg Arg Gly
    2000                 2005                 2010

Asp Leu  Pro Val Trp Leu Ala  Tyr Arg Val Ala Ala  Glu Gly Ile
    2015                 2020                 2025

Asn Tyr  Ala Asp Arg Arg Trp  Cys Phe Asp Gly Val  Lys Asn Asn
    2030                 2035                 2040

Gln Ile  Leu Glu Glu Asn Val  Glu Val Glu Ile Trp  Thr Lys Glu
    2045                 2050                 2055

Gly Glu  Arg Lys Lys Leu Lys  Pro Arg Trp Leu Asp  Ala Arg Ile
    2060                 2065                 2070
```

Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala
2075 2080 2085

Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met Gly Arg
2090 2095 2100

Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asn Ala Leu Asp Asn
2105 2110 2115

Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala Tyr Asn
2120 2125 2130

His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu Leu Leu
2135 2140 2145

Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu Phe Leu
2150 2155 2160

Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met Cys Cys
2165 2170 2175

Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile Gln Pro
2180 2185 2190

His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu Ile Val
2195 2200 2205

Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp Asn
2210 2215 2220

Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val Ala Ala
2225 2230 2235

Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp
2240 2245 2250

Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile
2255 2260 2265

Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala
2270 2275 2280

Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu
2285 2290 2295

Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala
2300 2305 2310

Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser Lys Met
2315 2320 2325

Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser Gln Val
2330 2335 2340

Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val Ala His
2345 2350 2355

Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr Arg Glu
2360 2365 2370

Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro Thr Val
2375 2380 2385

Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr Asp Pro
2390 2395 2400

Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val Leu Cys
2405 2410 2415

Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu Cys Glu
2420 2425 2430

Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly
2435 2440 2445

Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala
2450 2455 2460

Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe
2465 2470 2475

Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn
2480 2485 2490

Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala
2495 2500 2505

Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln
2510 2515 2520

Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu
2525 2530 2535

Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp
2540 2545 2550

Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val Val Asp
2555 2560 2565

Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly Gly Leu
2570 2575 2580

Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly Pro Gly
2585 2590 2595

His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn Leu Val
2600 2605 2610

Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu Lys
2615 2620 2625

Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro Asn Pro
2630 2635 2640

Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu Val Glu
2645 2650 2655

```
Asn Trp  Leu Asn Asn Asn Thr  Gln Phe Cys Ile Lys  Val Leu Asn
    2660                 2665                 2670

Pro Tyr  Met Pro Ser Val Ile  Glu Lys Met Glu Ala  Leu Gln Arg
    2675                 2680                 2685

Lys Tyr  Gly Gly Ala Leu Val  Arg Asn Pro Leu Ser  Arg Asn Ser
    2690                 2695                 2700

Thr His  Glu Met Tyr Trp Val  Ser Asn Ala Ser Gly  Asn Ile Val
    2705                 2710                 2715

Ser Ser  Val Asn Met Ile Ser  Arg Met Leu Ile Asn  Arg Phe Thr
    2720                 2725                 2730

Met Arg  Tyr Lys Lys Ala Thr  Tyr Glu Pro Asp Val  Asp Leu Gly
    2735                 2740                 2745

Ser Gly  Thr Arg Asn Ile Gly  Ile Glu Ser Glu Ile  Pro Asn Leu
    2750                 2755                 2760

Asp Ile  Ile Gly Lys Arg Ile  Glu Lys Ile Lys Gln  Glu His Glu
    2765                 2770                 2775

Thr Ser  Trp His Tyr Asp Gln  Asp His Pro Tyr Lys  Thr Trp Ala
    2780                 2785                 2790

Tyr His  Gly Ser Tyr Glu Thr  Lys Gln Thr Gly Ser  Ala Ser Ser
    2795                 2800                 2805

Met Val  Asn Gly Val Val Arg  Leu Leu Thr Lys Pro  Trp Asp Val
    2810                 2815                 2820

Val Pro  Met Val Thr Gln Met  Ala Met Thr Asp Thr  Thr Pro Phe
    2825                 2830                 2835

Gly Gln  Gln Arg Val Phe Lys  Glu Lys Val Asp Thr  Arg Thr Gln
    2840                 2845                 2850
```

Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr Ala Glu
2855 2860 2865

Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg Met Cys
2870 2875 2880

Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala Ala Leu
2885 2890 2895

Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala Arg Glu
2900 2905 2910

Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys Glu Arg
2915 2920 2925

Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn Met
2930 2935 2940

Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
2945 2950 2955

Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu
2960 2965 2970

Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser
2975 2980 2985

Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys
2990 2995 3000

Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly Gly Ala
3005 3010 3015

Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Leu
3020 3025 3030

Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met Glu Gly
3035 3040 3045

```
Glu His  Lys Lys Leu Ala Glu  Ala Ile Phe Lys Leu  Thr Tyr Gln
    3050                 3055                 3060

Asn Lys  Val Val Arg Val Gln  Arg Pro Thr Pro Arg  Gly Thr Val
    3065                 3070                 3075

Met Asp  Ile Ile Ser Arg Arg  Asp Gln Arg Gly Ser  Gly Gln Val
    3080                 3085                 3090

Gly Thr  Tyr Gly Leu Asn Thr  Phe Thr Asn Met Glu  Ala Gln Leu
    3095                 3100                 3105

Ile Arg  Gln Met Glu Gly Glu  Gly Val Phe Lys Ser  Ile Gln His
    3110                 3115                 3120

Leu Thr  Ile Thr Glu Glu Ile  Ala Val Gln Asn Trp  Leu Ala Arg
    3125                 3130                 3135

Val Gly  Arg Glu Arg Leu Ser  Arg Met Ala Ile Ser  Gly Asp Asp
    3140                 3145                 3150

Cys Val  Val Lys Pro Leu Asp  Asp Arg Phe Ala Ser  Ala Leu Thr
    3155                 3160                 3165

Ala Leu  Asn Asp Met Gly Lys  Ile Arg Lys Asp Ile  Gln Gln Trp
    3170                 3175                 3180

Glu Pro  Ser Arg Gly Trp Asn  Asp Trp Thr Gln Val  Pro Phe Cys
    3185                 3190                 3195

Ser His  His Phe His Glu Leu  Ile Met Lys Asp Gly  Arg Val Leu
    3200                 3205                 3210

Val Val  Pro Cys Arg Asn Gln  Asp Glu Leu Ile Gly  Arg Ala Arg
    3215                 3220                 3225

Ile Ser  Gln Gly Ala Gly Trp  Ser Leu Arg Glu Thr  Ala Cys Leu
    3230                 3235                 3240
```

Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg
3245 3250 3255

Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro
3260 3265 3270

Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile His Ala
3275 3280 3285

Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val Trp Asn
3290 3295 3300

Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys Thr Pro
3305 3310 3315

Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu Asp
3320 3325 3330

Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp
3335 3340 3345

Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile
3350 3355 3360

Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe
3365 3370 3375

Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385

<210> 23
<211> 10717
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 3, BVS

<400> 23
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgatg gagagccgcg catgattgtg 480

gggaagaatg aaagaggaaa atccctactt ttcaagacag cctctggaat caacatgtgc 540

acactcatag ctatggatct gggagagatg tgtgatgaca cggtcactta caaatgcccc 600

cacattaccg aagtggagcc tgaagacatt gactgctggt gcaaccttac atcgacatgg 660

gtgacttatg gaacatgcaa tcaagctgga gagcatagac gcgataagag atcagtggcg 720

ttagctcccc atgttggcat gggactggac acacgcactc aaacctggat gtcggctgaa 780

ggagcttgga gacaagtcga gaaggtagag acatgggccc ttaggcaccc agggtttacc 840

atactagccc tatttcttgc ccattacata ggcacttcct tgacccagaa agtggttatt 900

tttatactat taatgctggt taccccatcc atgacaatga gatgtgtagg agtaggaaac 960

agagattttg tggaaggcct atcgggagct acgtgggttg acgtggtgct cgagcacggt 1020

gggtgtgtga ctaccatggc taagaacaag cccacgctgg acatagagct tcagaagacc 1080

gaggccaccc aactggcgac cctaaggaag ctatgcattg agggaaaaat taccaacata 1140

acaaccgact caagatgtcc cacccaaggg gaagcgattt tacctgagga gcaggaccag 1200

aactacgtgt gtaagcatac atacgtggac agaggctggg gaaacggttg tggtttgttt 1260

ggcaagggaa gcttggtgac atgcgcgaaa tttcaatgtt tagaatcaat agagggaaaa 1320

gtggtgcaac atgagaacct caaatacacc gtcatcatca cagtgcacac aggagaccaa 1380

caccaggtgg gaaatgaaac gcagggagtc acggctgaga taacacccca ggcatcaacc 1440

gctgaagcca ttttacctga atatggaacc ctcgggctag aatgctcacc acggacaggt 1500

ttggatttca atgaaatgat ctcattgaca atgaagaaca aagcatggat ggtacataga 1560

caatggttct ttgacttacc cctaccatgg acatcaggag cttcagcaga aacaccaact 1620

tggaacagga aagagcttct tgtgacattt aaaaatgcac atgcaaaaaa gcaagaagta 1680

gttgttcttg gatcacaaga gggagcaatg catacagcac tgacaggagc tacagagatc 1740

caaacctcag gaggcacaag tatctttgcg gggcacttaa aatgtagact caagatggac 1800

aaattggaac tcaaggggat gagctatgca atgtgcttga gtagctttgt gttgaagaaa 1860

gaagtctccg aaacgcagca tgggacaata ctcattaagg ttgagtacaa aggggaagat 1920

gcaccctgca agattccttt ctccacggag gatggacaag gaaaagctct caatggcaga 1980

ctgatcacag ccaatccagt ggtgaccaag aaggaggagc ctgtcaacat tgaggctgaa 2040

cctccttttg gagaaagtaa catagtaatt ggaattggag acaaagccct gaaaatcaac 2100

tggtacaaga agggaagctc gattgggaag atgttcgagg ccactgccag aggtgcaagg 2160

cgcatggcca tcttgggaga cacagcctgg gactttggat cagtgggtgg tgttttgaat 2220

tcattaggga aaatggtcca ccaaatattt gggagtgctt acacagccct atttggtgga 2280

gtctcctgga tgatgaaaat tggaataggt gtcctcttaa cctggatagg gttgaactca 2340

aaaaatactt ctatgtcatt ttcatgcatc gcggccggca ttgtgacact gtatttggga 2400

gtcatggtgc aggccgatag tggttgcgtt gtgagctgga aaaacaaaga actgaaatgt 2460

ggcagtggga ttttcatcac agacaacgtg cacacatgga cagaacaata caagttccaa 2520

ccagaatccc cttcaaaact agcttcagct atccagaaag cccatgaaga ggacatttgt 2580

ggaatccgct cagtaacaag actggagaat ctgatgtgga aacaaataac accagaattg 2640

aatcacattc tatcagaaaa tgaggtgaag ttaactatta tgacaggaga catcaaagga 2700

atcatgcagg caggaaaacg atctctgcgg cctcagccca ctgagctgaa gtattcatgg 2760

aaaacatggg gcaaagcaaa aatgctctct acagagtctc ataaccagac ctttctcatt 2820

gatggccccg aaacagcaga atgccccaac acaaatagag cttggaattc gttggaagtt 2880

gaagactatg gctttggagt attcaccacc aatatatggc taaaattgaa agaaaaacag 2940

gatgtattct gcgactcaaa actcatgtca gcggccataa aagacaacag agccgtccat 3000

gccgatatgg gttattggat agaaagtgca ctcaatgaca catggaagat agagaaagcc 3060

tctttcattg aagttaaaaa ctgccactgg ccaaaatcac acaccctctg gagcaatgga 3120

gtgctagaaa gtgagatgat aattccaaag aatctcgctg gaccagtgtc tcaacacaac 3180

tatagaccag gctaccatac acaaataaca ggaccatggc atctaggtaa gcttgagatg 3240

gactttgatt tctgtgatgg aacaacagtg gtagtgactg aggactgcgg aaatagagga 3300

ccctctttga gaacaaccac tgcctctgga aaactcataa cagaatggtg ctgccgatct 3360

tgcacattac caccgctaag atacagaggt gaggatgggt gctggtacgg gatggaaatc 3420

agaccattga aggagaaaga agagaatttg gtcaactcct tggtcacagc tggacatggg 3480

caggtcgaca acttttcact aggagtcttg ggaatggcat tgttcctgga ggaaatgctt 3540

aggacccgag taggaacgaa acatgcaata ctactagttg cagtttcttt tgtgacattg 3600

atcacaggga acatgtcctt tagagacctg ggaagagtga tggttatggt aggcgccact 3660

atgacggatg acataggtat gggcgtgact tatcttgccc tactagcagc cttcaaagtc 3720

agaccaactt ttgcagctgg actactcttg agaaagctga cctccaagga attgatgatg 3780

actactatag gaattgtact cctctcccag agcaccatac cagagaccat tcttgagttg 3840

actgatgcgt tagccttagg catgatggtc ctcaaaatgg tgagaaatat ggaaaagtat 3900

caattggcag tgactatcat ggctatcttg tgcgtcccaa acgcagtgat attacaaaac 3960

gcatggaaag tgagttgcac aatattggca gtggtgtccg tttccccact gttcttaaca 4020

tcctcacagc aaaaaacaga ttggatacca ttagcattga cgatcaaagg tctcaatcca 4080

acagctattt ttctaacaac cctctcaaga accagcaaga aaaggagctg gccattaaat 4140

gaggctatca tggcagtcgg gatggtgagc attttagcca gttctctcct aaaaaatgat 4200

attcccatga caggaccatt agtggctgga gggctcctca ctgtgtgcta cgtgctcact 4260

ggacgatcgg ccgatttgga actggagaga gcagccgatg tcaaatggga agaccaggca 4320

gagatatcag gaagcagtcc aatcctgtca ataacaatat cagaagatgg tagcatgtcg 4380

ataaaaaatg aagaggaaga acaaacactg accatactca ttagaacagg attgctggtg 4440

atctcaggac tttttcctgt atcaatacca atcacggcag cagcatggta cctgtgggaa 4500

gtgaagaaac aacgggccgg agtattgtgg gatgttcctt cacccccacc catgggaaag 4560

gctgaactgg aagatggagc ctatagaatt aagcaaaaag ggattcttgg atattcccag 4620

atcggagccg gagtttacaa agaaggaaca ttccatacaa tgtggcatgt cacacgtggc 4680

gctgttctaa tgcataaagg aaagaggatt gaaccatcat gggcggacgt caagaaagac 4740

ctaatatcat atggaggagg ctggaagtta gaaggagaat ggaaggaagg agaagaagtc 4800

caggtattgg cactggagcc tggaaaaaat ccaagagccg tccaaacgaa acctggtctt 4860

ttcaaaacca acgccggaac aataggtgct gtatctctgg acttttctcc tggaacgtca 4920

ggatctccaa ttatcgacaa aaaaggaaaa gttgtgggtc tttatggtaa tggtgttgtt 4980

acaaggagtg gagcatatgt gagtgctata gcccagactg aaaaaagcat tgaagacaac 5040

ccagagatcg aagatgacat tttccgaaag agaagactga ccatcatgga cctccaccca 5100

ggagcgggaa agacgaagag ataccttccg gccatagtca gagaagctat aaaacggggt 5160

ttgagaacat taatcttggc ccccactaga gttgtggcag ctgaaatgga ggaagccctt 5220

agaggacttc caataagata ccagacccca gccatcagag ctgtgcacac cgggcgggag 5280

attgtggacc taatgtgtca tgccacattt accatgaggc tgctatcacc agttagagtg 5340

ccaaactaca acctgattat catggacgaa gcccatttca cagacccagc aagtatagca 5400

gctagaggat acatctcaac tcgagtggag atgggtgagg cagctgggat ttttatgaca 5460

gccactcccc cgggaagcag agacccattt cctcagagca atgcaccaat catagatgaa 5520

gaaagagaaa tccctgaacg ctcgtggaat tccggacatg aatgggtcac ggattttaaa 5580

gggaagactg tttggttcgt tccaagtata aaagcaggaa atgatatagc agcttgcctg 5640

aggaaaaatg gaaagaaagt gatacaactc agtaggaaga cctttgattc tgagtatgtc 5700

aagactagaa ccaatgattg ggacttcgtg gttacaactg acatttcaga aatgggtgcc 5760

aatttcaagg ctgagagggt tatagacccc agacgctgca tgaaaccagt catactaaca 5820

gatggtgaag agcgggtgat tctggcagga cctatgccag tgacccactc tagtgcagca 5880

caaagaagag ggagaatagg aagaaatcca aaaaatgaga atgaccagta catatacatg 5940

ggggaacctc tggaaaatga tgaagactgt gcacactgga aagaagctaa aatgctccta 6000

gataacatca acacgccaga aggaatcatt cctagcatgt tcgaaccaga gcgtgaaaag 6060

gtggatgcca ttgatggcga ataccgcttg agaggagaag caaggaaaac ctttgtagac 6120

ttaatgagaa gaggagacct accagtctgg ttggcctaca gagtggcagc tgaaggcatc 6180

aactacgcag acagaaggtg gtgttttgat ggagtcaaga acaaccaaat cctagaagaa 6240

aacgtggaag ttgaaatctg gacaaaagaa ggggaaagga agaaattgaa acccagatgg 6300

ttggatgcta ggatctattc tgacccactg gcgctaaaag aatttaagga atttgcagcc 6360

ggaagaaagt ctctgaccct gaacctaatc acagaaatgg gtaggctccc aaccttcatg 6420

actcagaagg caagaaacgc actggacaac ttagcagtgc tgcacacggc tgaggcaggt 6480

ggaagggcgt acaaccatgc tctcagtgaa ctgccggaga ccctggagac attgctttta 6540

ctgacacttc tggctacagt cacgggaggg atctttttat tcttgatgag cgcaaggggc 6600

atagggaaga tgaccctggg aatgtgctgc ataatcacgg ctagcatcct cctatggtac 6660

gcacaaatac agccacactg gatagcagct tcaataatac tggagttttt tctcatagtt 6720

ttgcttattc cagaacctga aaaacagaga acaccccaag acaaccaact gacctacgtt 6780

gtcatagcca tcctcacagt ggtggccgca accatggcaa acgagatggg tttcctagaa 6840

aaaacgaaga aagatctcgg attgggaagc attgcaaccc agcaacccga gagcaacatc 6900

ctggacatag atctacgtcc tgcatcagca tggacgctgt atgccgtggc cacaacattt 6960

gttacaccaa tgttgagaca tagcattgaa aattcctcag tgaatgtgtc cctaacagct 7020

atagccaacc aagccacagt gttaatgggt ctcgggaaag gatggccatt gtcaaagatg 7080

gacatcggag ttccccttct cgccattgga tgctactcac aagtcaaccc cataactctc 7140

acagcagctc ttttcttatt ggtagcacat tatgccatca tagggccagg actccaagca 7200

aaagcaacca gagaagctca gaaaagagca gcggcgggca tcatgaaaaa cccaactgtc 7260

gatggaataa cagtgattga cctagatcca ataccttatg atccaaagtt tgaaaagcag 7320

ttgggacaag taatgctcct agtcctctgc gtgactcaag tattgatgat gaggactaca 7380

tgggctctgt gtgaggcttt aaccttagct accgggccca tctccacatt gtgggaagga 7440

aatccaggga ggttttggaa cactaccatt gcggtgtcaa tggctaacat ttttagaggg 7500

agttacttgg ccggagctgg acttctcttt tctattatga agaacacaac caacacaaga 7560

aggggaactg gcaacatagg agagacgctt ggagagaaat ggaaaagccg attgaacgcg 7620

ttgggaaaaa gtgaattcca gatctacaag aaaagtggaa tccaggaagt ggatagaacc 7680

ttagcaaaag aaggcattaa aagaggagaa acggaccatc acgctgtgtc gcgaggctca 7740

gcaaaactga gatggttcgt tgagagaaac atggtcacac cagaagggaa agtagtggac 7800

ctcggttgtg gcagaggagg ctggtcatac tattgtggag gactaaagaa tgtaagagaa 7860

gtcaaaggcc taacaaaagg aggaccagga cacgaagaac ccatccccat gtcaacatat 7920

gggtggaatc tagtgcgtct tcaaagtgga gttgacgttt tcttcatccc gccagaaaag 7980

tgtgacacat tattgtgtga catagggggag tcatcaccaa atcccacagt ggaagcagga 8040

cgaacactca gagtccttaa cttagtagaa aattggttga acaacaacac tcaattttgc 8100

ataaaggttc tcaacccata tatgccctca gtcatagaaa aaatggaagc actacaaagg 8160

aaatatggag gagccttagt gaggaatcca ctctcacgaa actccacaca tgagatgtac 8220

tgggtatcca atgcttccgg gaacatagtg tcatcagtga acatgatttc aaggatgttg 8280

atcaacagat ttacaatgag atacaagaaa gccacttacg agccggatgt tgacctcgga 8340

agcggaaccc gtaacatcgg gattgaaagt gagataccaa acctagatat aattgggaaa 8400

agaatagaaa aaataaagca agagcatgaa acatcatggc actatgacca agaccaccca 8460

tacaaaacgt gggcatacca tggtagctat gaaacaaaac agactggatc agcatcatcc 8520

atggtcaacg gagtggtcag gctgctgaca aaaccttggg acgtcgtccc catggtgaca 8580

cagatggcaa tgacagacac gactccattt ggacaacagc gcgtttttaa agagaaagtg 8640

gacacgagaa cccaagaacc gaaagaaggc acgaagaaac taatgaaaat aacagcagag 8700

tggctttgga aagaattagg gaagaaaaag acacccagga tgtgcaccag agaagaattc 8760

acaagaaagg tgagaagcaa tgcagccttg ggggccatat tcactgatga gaacaagtgg 8820

aagtcggcac gtgaggctgt tgaagatagt aggttttggg agctggttga caaggaaagg 8880

aatctccatc ttgaaggaaa gtgtgaaaca tgtgtgtaca acatgatggg aaaaagagag 8940

aagaagctag gggaattcgg caaggcaaaa ggcagcagag ccatatggta catgtggctt 9000

ggagcacgct tcttagagtt tgaagcccta ggattcttaa atgaagatca ctggttctcc 9060

agagagaact ccctgagtgg agtggaagga gaagggctgc acaagctagg ttacattcta 9120

agagacgtga gcaagaaaga gggaggagca atgtatgccg atgacaccgc aggatgggat 9180

acaagaatca cactagaaga cctaaaaaat gaagaaatgg taacaaacca catggaagga 9240

gaacacaaga aactagccga ggccattttc aaactaacgt accaaaacaa ggtggtgcgt 9300

gtgcaaagac caacaccaag aggcacagta atggacatca tatcgagaag agaccaaaga 9360

ggtagtggac aagttggcac ctatggactc aatactttca ccaatatgga agcccaacta 9420

atcagacaga tggagggaga aggagtcttt aaaagcattc agcacctaac aatcacagaa 9480

gaaatcgctg tgcaaaactg gttagcaaga gtggggcgcg aaaggttatc aagaatggcc 9540

atcagtggag atgattgtgt tgtgaaacct ttagatgaca ggttcgcaag cgctttaaca 9600

gctctaaatg acatgggaaa gattaggaaa gacatacaac aatgggaacc ttcaagagga 9660

tggaatgatt ggacacaagt gcccttctgt tcacaccatt tccatgagtt aatcatgaaa 9720

gacggtcgcg tactcgttgt tccatgtaga aaccaagatg aactgattgg cagagcccga 9780

atctcccaag gagcagggtg gtctttgcgg gagacggcct gtttggggaa gtcttacgcc 9840

caaatgtgga gcttgatgta cttccacaga cgcgacctca ggctggcggc aaatgctatt 9900

tgctcggcag taccatcaca ttgggttcca acaagtcgaa caacctggtc catacatgct 9960

aaacatgaat ggatgacaac ggaagacatg ctgacagtct ggaacagggt gtggattcaa 10020

gaaaacccat ggatggaaga caaaactcca gtggaatcat gggaggaaat cccatacttg 10080

gggaaaagag aagaccaatg gtgcggctca ttgattgggt taacaagcag ggccacctgg 10140

gcaaagaaca tccaagcagc aataaatcaa gttagatccc ttataggcaa tgaagaatac 10200

acagattaca tgccatccat gaaaagattc agaagagaag aggaagaagc aggagttctg 10260

tggtagaaag caaaactaac atgaaacaag gctagaagtc aggtcggatt aagccatagt 10320

acggaaaaaa ctatgctacc tgtgagcccc gtccaaggac gttaaaagaa gtcaggccat 10380

cataaatgcc atagcttgag taaactatgc agcctgtagc tccacctgag aaggtgtaaa 10440

aaatccggga ggccacaaac catggaagct gtacgcatgg cgtagtggac tagcggttag 10500

aggagacccc tcccttacaa atcgcagcaa caatgggggc ccaaggcgag atgaagctgt 10560

agtctcgctg gaaggactag aggttagagg agacccccccc gaaacaaaaa acagcatatt 10620

gacgctggga aagaccagag atcctgctgt ctcctcagca tcattccagg cacagaacgc 10680

cagaaaatgg aatggtgctg ttgaatcaac aggttct 10717

<210> 24
<211> 3389
<212> PRT
<213> Trình tự nhân tạo

<220>

<223> Typ huyết thanh virut sốt xuất huyết 3, BVS

<400> 24

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Arg Met Ile Val
115 120 125

Gly Lys Asn Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ala Ser Gly
130 135 140

Ile Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Asp
145 150 155 160

Asp Thr Val Thr Tyr Lys Cys Pro His Ile Thr Glu Val Glu Pro Glu
165 170 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Thr Tyr Gly
180 185 190

```
Thr Cys Asn Gln Ala Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
        195                 200                 205

Leu Ala Pro His Val Gly Met Gly Leu Asp Thr Arg Thr Gln Thr Trp
    210                 215                 220

Met Ser Ala Glu Gly Ala Trp Arg Gln Val Glu Lys Val Glu Thr Trp
225                 230                 235                 240

Ala Leu Arg His Pro Gly Phe Thr Ile Leu Ala Leu Phe Leu Ala His
                245                 250                 255

Tyr Ile Gly Thr Ser Leu Thr Gln Lys Val Val Ile Phe Ile Leu Leu
            260                 265                 270

Met Leu Val Thr Pro Ser Met Thr Met Arg Cys Val Gly Val Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
    290                 295                 300

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Ile Glu Leu Gln Lys Thr Glu Ala Thr Gln Leu Ala Thr Leu
                325                 330                 335

Arg Lys Leu Cys Ile Glu Gly Lys Ile Thr Asn Ile Thr Thr Asp Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Ala Ile Leu Pro Glu Glu Gln Asp Gln
        355                 360                 365

Asn Tyr Val Cys Lys His Thr Tyr Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Val Thr Cys Ala Lys Phe Gln
385                 390                 395                 400
```

Cys Leu Glu Ser Ile Glu Gly Lys Val Val Gln His Glu Asn Leu Lys
405 410 415

Tyr Thr Val Ile Ile Thr Val His Thr Gly Asp Gln His Gln Val Gly
420 425 430

Asn Glu Thr Gln Gly Val Thr Ala Glu Ile Thr Pro Gln Ala Ser Thr
435 440 445

Ala Glu Ala Ile Leu Pro Glu Tyr Gly Thr Leu Gly Leu Glu Cys Ser
450 455 460

Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Ile Ser Leu Thr Met Lys
465 470 475 480

Asn Lys Ala Trp Met Val His Arg Gln Trp Phe Phe Asp Leu Pro Leu
485 490 495

Pro Trp Thr Ser Gly Ala Ser Ala Glu Thr Pro Thr Trp Asn Arg Lys
500 505 510

Glu Leu Leu Val Thr Phe Lys Asn Ala His Ala Lys Lys Gln Glu Val
515 520 525

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

Ala Thr Glu Ile Gln Thr Ser Gly Gly Thr Ser Ile Phe Ala Gly His
545 550 555 560

Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Glu Leu Lys Gly Met Ser
565 570 575

Tyr Ala Met Cys Leu Ser Ser Phe Val Leu Lys Lys Glu Val Ser Glu
580 585 590

Thr Gln His Gly Thr Ile Leu Ile Lys Val Glu Tyr Lys Gly Glu Asp
595 600 605

```
Ala Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp Gly Gln Gly Lys Ala
    610                 615                 620

Leu Asn Gly Arg Leu Ile Thr Ala Asn Pro Val Val Thr Lys Lys Glu
625                 630                 635                 640

Glu Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser Asn Ile
                645                 650                 655

Val Ile Gly Ile Gly Asp Lys Ala Leu Lys Ile Asn Trp Tyr Lys Lys
            660                 665                 670

Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly Ala Arg
        675                 680                 685

Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly
    690                 695                 700

Gly Val Leu Asn Ser Leu Gly Lys Met Val His Gln Ile Phe Gly Ser
705                 710                 715                 720

Ala Tyr Thr Ala Leu Phe Gly Gly Val Ser Trp Met Met Lys Ile Gly
                725                 730                 735

Ile Gly Val Leu Leu Thr Trp Ile Gly Leu Asn Ser Lys Asn Thr Ser
            740                 745                 750

Met Ser Phe Ser Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr Leu Gly
        755                 760                 765

Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys Asn Lys
    770                 775                 780

Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val His Thr
785                 790                 795                 800

Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys Leu Ala
                805                 810                 815
```

Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile Arg Ser
820 825 830

Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro Glu Leu
835 840 845

Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met Thr Gly
850 855 860

Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg Pro Gln
865 870 875 880

Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala Lys Met
885 890 895

Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly Pro Glu
900 905 910

Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu Glu Val
915 920 925

Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu Lys Leu
930 935 940

Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser Ala Ala
945 950 955 960

Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp Ile Glu
965 970 975

Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe Ile Glu
980 985 990

Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser Asn Gly
995 1000 1005

Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala Gly Pro
1010 1015 1020

Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln Ile Thr
1025 1030 1035

Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp Phe Cys
1040 1045 1050

Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn Arg Gly
1055 1060 1065

Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile Thr Glu
1070 1075 1080

Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr Arg Gly
1085 1090 1095

Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu Lys Glu
1100 1105 1110

Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly His Gly
1115 1120 1125

Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala Leu Phe
1130 1135 1140

Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His Ala Ile
1145 1150 1155

Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly Asn Met
1160 1165 1170

Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly Ala Thr
1175 1180 1185

Met Thr Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala Leu Leu
1190 1195 1200

Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu Leu Leu
1205 1210 1215

```
Arg Lys  Leu Thr Ser Lys Glu  Leu Met Met Thr Thr  Ile Gly Ile
    1220                 1225                 1230

Val Leu  Leu Ser Gln Ser Thr  Ile Pro Glu Thr Ile  Leu Glu Leu
    1235                 1240                 1245

Thr Asp  Ala Leu Ala Leu Gly  Met Met Val Leu Lys  Met Val Arg
    1250                 1255                 1260

Asn Met  Glu Lys Tyr Gln Leu  Ala Val Thr Ile Met  Ala Ile Leu
    1265                 1270                 1275

Cys Val  Pro Asn Ala Val Ile  Leu Gln Asn Ala Trp  Lys Val Ser
    1280                 1285                 1290

Cys Thr  Ile Leu Ala Val Val  Ser Val Ser Pro Leu  Phe Leu Thr
    1295                 1300                 1305

Ser Ser  Gln Gln Lys Thr Asp  Trp Ile Pro Leu Ala  Leu Thr Ile
    1310                 1315                 1320

Lys Gly  Leu Asn Pro Thr Ala  Ile Phe Leu Thr Thr  Leu Ser Arg
    1325                 1330                 1335

Thr Ser  Lys Lys Arg Ser Trp  Pro Leu Asn Glu Ala  Ile Met Ala
    1340                 1345                 1350

Val Gly  Met Val Ser Ile Leu  Ala Ser Ser Leu Leu  Lys Asn Asp
    1355                 1360                 1365

Ile Pro  Met Thr Gly Pro Leu  Val Ala Gly Gly Leu  Leu Thr Val
    1370                 1375                 1380

Cys Tyr  Val Leu Thr Gly Arg  Ser Ala Asp Leu Glu  Leu Glu Arg
    1385                 1390                 1395

Ala Ala  Asp Val Lys Trp Glu  Asp Gln Ala Glu Ile  Ser Gly Ser
    1400                 1405                 1410
```

Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser Met Ser
1415 1420 1425

Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu Ile Arg
1430 1435 1440

Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser Ile Pro
1445 1450 1455

Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys Gln Arg
1460 1465 1470

Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met Gly Lys
1475 1480 1485

Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys Gly Ile
1490 1495 1500

Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu Gly Thr
1505 1510 1515

Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu Met His
1520 1525 1530

Lys Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys Lys Asp
1535 1540 1545

Leu Ile Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu Trp Lys
1550 1555 1560

Glu Gly Glu Glu Val Gln Val Leu Ala Leu Glu Pro Gly Lys Asn
1565 1570 1575

Pro Arg Ala Val Gln Thr Lys Pro Gly Leu Phe Lys Thr Asn Ala
1580 1585 1590

Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser Pro Gly Thr Ser
1595 1600 1605

Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly Leu Tyr
1610 1615 1620

Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser Ala Ile
1625 1630 1635

Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile Glu Asp
1640 1645 1650

Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu His Pro
1655 1660 1665

Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val Arg Glu
1670 1675 1680

Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro Thr Arg
1685 1690 1695

Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu Pro Ile
1700 1705 1710

Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly Arg Glu
1715 1720 1725

Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg Leu Leu
1730 1735 1740

Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met Asp Glu
1745 1750 1755

Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly Tyr Ile
1760 1765 1770

Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe Met Thr
1775 1780 1785

Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser Asn Ala
1790 1795 1800

Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser Trp Asn
1805 1810 1815

Ser Gly His Glu Trp Val Thr Asp Phe Lys Gly Lys Thr Val Trp
1820 1825 1830

Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala Cys Leu
1835 1840 1845

Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg Lys Thr Phe
1850 1855 1860

Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp Phe Val
1865 1870 1875

Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys Ala Glu
1880 1885 1890

Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile Leu Thr
1895 1900 1905

Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro Val Thr
1910 1915 1920

His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg Asn Pro
1925 1930 1935

Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro Leu Glu
1940 1945 1950

Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met Leu Leu
1955 1960 1965

Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met Phe Glu
1970 1975 1980

Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr Arg Leu
1985 1990 1995

Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg Arg Gly
2000 2005 2010

Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu Gly Ile
2015 2020 2025

Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys Asn Asn
2030 2035 2040

Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr Lys Glu
2045 2050 2055

Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala Arg Ile
2060 2065 2070

Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe Ala Ala
2075 2080 2085

Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met Gly Arg
2090 2095 2100

Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asn Ala Leu Asp Asn
2105 2110 2115

Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala Tyr Asn
2120 2125 2130

His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu Leu Leu
2135 2140 2145

Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu Phe Leu
2150 2155 2160

Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met Cys Cys
2165 2170 2175

Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile Gln Pro
2180 2185 2190

His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu Ile Val
2195 2200 2205

Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln Asp Asn
2210 2215 2220

Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val Ala Ala
2225 2230 2235

Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys Lys Asp
2240 2245 2250

Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser Asn Ile
2255 2260 2265

Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu Tyr Ala
2270 2275 2280

Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser Ile Glu
2285 2290 2295

Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn Gln Ala
2300 2305 2310

Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser Lys Met
2315 2320 2325

Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser Gln Val
2330 2335 2340

Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val Ala His
2345 2350 2355

Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr Arg Glu
2360 2365 2370

Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro Thr Val
2375 2380 2385

Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr Asp Pro
2390 2395 2400

Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val Leu Cys
2405 2410 2415

Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu Cys Glu
2420 2425 2430

Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp Glu Gly
2435 2440 2445

Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser Met Ala
2450 2455 2460

Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu Leu Phe
2465 2470 2475

Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr Gly Asn
2480 2485 2490

Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu Asn Ala
2495 2500 2505

Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly Ile Gln
2510 2515 2520

Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg Gly Glu
2525 2530 2535

Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu Arg Trp
2540 2545 2550

Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val Val Asp
2555 2560 2565

Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly Gly Leu
2570 2575 2580

Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly Pro Gly
2585 2590 2595

His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn Leu Val
2600 2605 2610

Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro Glu Lys
2615 2620 2625

Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro Asn Pro
2630 2635 2640

Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu Val Glu
2645 2650 2655

Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val Leu Asn
2660 2665 2670

Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu Gln Arg
2675 2680 2685

Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg Asn Ser
2690 2695 2700

Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn Ile Val
2705 2710 2715

Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg Phe Thr
2720 2725 2730

Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp Leu Gly
2735 2740 2745

Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro Asn Leu
2750 2755 2760

Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu His Glu
2765 2770 2775

Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr Trp Ala
2780 2785 2790

Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala Ser Ser
2795 2800 2805

Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp Asp Val
2810 2815 2820

Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr Pro Phe
2825 2830 2835

Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg Thr Gln
2840 2845 2850

Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr Ala Glu
2855 2860 2865

Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg Met Cys
2870 2875 2880

Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala Ala Leu
2885 2890 2895

Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala Arg Glu
2900 2905 2910

Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys Glu Arg
2915 2920 2925

Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr Asn Met
2930 2935 2940

Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys Ala Lys
2945 2950 2955

Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg Phe Leu
2960 2965 2970

Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp Phe Ser
2975 2980 2985

Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu His Lys
2990 2995 3000

Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly Gly Ala
3005 3010 3015

Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile Thr Leu
3020 3025 3030

Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met Glu Gly
3035 3040 3045

Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr Tyr Gln
3050 3055 3060

Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly Thr Val
3065 3070 3075

Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly Gln Val
3080 3085 3090

Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala Gln Leu
3095 3100 3105

Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile Gln His
3110 3115 3120

Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu Ala Arg
3125 3130 3135

Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly Asp Asp
3140 3145 3150

Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala Leu Thr
3155 3160 3165

Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln Gln Trp
3170 3175 3180

Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro Phe Cys
3185 3190 3195

Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg Val Leu
3200 3205 3210

Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg Ala Arg
3215 3220 3225

Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala Cys Leu
3230 3235 3240

Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe His Arg
3245 3250 3255

Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala Val Pro
3260 3265 3270

Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile His Ala
3275 3280 3285

Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val Trp Asn
3290 3295 3300

Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys Thr Pro
3305 3310 3315

Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg Glu Asp
3320 3325 3330

Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala Thr Trp
3335 3340 3345

Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser Leu Ile
3350 3355 3360

Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys Arg Phe
3365 3370 3375

Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385

<210> 25
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 4, Tiền chủng gốc

<400> 25
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gacctttaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagctctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgatg gcgaacccct catgatagtg 480

gcaaaacatg aaaggggggag acctctcttg tttaagacaa cagaggggat caacaaatgc 540

actctcattg ccatggactt gggtgaaatg tgtgaggaca ctgtcacgta taaatgcccc 600

ttactggtca ataccgaacc tgaagacatt gattgctggt gcaatctcac gtctacctgg 660

gtcatgtatg ggacatgcac ccagagcgga gaacggagac gagagaagcg ctcagtagct 720

ttaacaccac attcaggaat gggattggaa acaagagctg agacatggat gtcatcggaa 780

ggggcttgga agcatgctca gagagtagag agctggatac tcagaaaccc aggattcgcg 840

ctcttggcag gatttatggc ttatatgatt gggcaaacag gaatccagcg aactgtcttc 900

tttgtcctaa tgatgctggt cgccccatcc tacggaatgc gatgcgtagg agtaggaaac 960

agagactttg tggaaggagt ctcaggtgga gcatgggtcg atctggtgct agaacatgga 1020

ggatgcgtca caaccatggc ccagggaaaa ccaaccttgg attttgaact gactaagaca 1080

acagccaagg aagtggctct gttaagaacc tattgcattg aagcctcaat atcaaacata 1140

accacggcaa caagatgtcc aacgcaagga gagccttatc taaaagagga acaagaccaa 1200

cagtacattt gccggagaga tgtggtagac agagggtggg gcaatggctg tggcttgttt 1260

ggaaaaggag gagttgtgac atgtgcgaag ttttcatgtt cggggaagat aacaggcaat 1320

ttggtccaaa ttgagaacct tgaatacaca gtggttgtaa cagtccacaa tggagacacc 1380

catgcagtag gaaatgacac gtccaatcat ggagttacag ccacgataac tcccaggtca 1440

ccatcggtgg aagtcaaatt gccggactat ggagaactaa cactcgattg tgaacccagg 1500

tctggaattg actttaatga gatgattctg atgaaaatga aaaagaaaac atggcttgtg 1560

cataagcaat ggtttttgga tctacctcta ccatggacag caggagcaga cacatcagag 1620

gttcactgga attacaaaga gagaatggtg acatttaagg ttcctcatgc caagagacag 1680

gatgtgacag tgctgggatc tcaggaagga gccatgcatt ctgccctcgc tggagccaca 1740

gaagtggact ccggtgatgg aaatcacatg tttgcaggac atctcaagtg caaagtccgt 1800

atggagaaat tgagaatcaa gggaatgtca tacacgatgt gttcaggaaa gttctcaatt 1860

gacaaagaga tggcagaaac acagcatggg acaacagtgg tgaaagtcaa gtatgaaggt 1920

gctggagctc cgtgtaaagt ccccatagag ataagagatg tgaacaagga aaaagtggtt 1980

gggcgtatca tctcatccac ccctttggct gagaatacca acagtgtaac caacatagag 2040

ttagaacccc cctttgggga cagctacata gtgataggtg ttggaaacag tgcattaaca 2100

ctccattggt tcaggaaagg gagttccatt ggcaagatgt ttgagtccac atacagaggt 2160

gcaaaacgaa tggccattct aggtgaaaca gcttgggatt ttggttccgt tggtggactg 2220

ttcacatcat tgggaaaggc tgtgcaccag gtttttggaa gtgtgtatac aaccctgttt 2280

ggaggagtct catggatgat tagaatccta attgggttcc tagtgttgtg gattggcacg 2340

aactcaagga acacttcaat ggctatgacg tgcatagctg ccggcattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cgggtgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga tccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggtaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gtttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagccatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttccc tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 26
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 4, Tiền chủng gốc

<400> 26

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

Arg Arg Arg Ser Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
100 105 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Leu Met Ile Val
115 120 125

Ala Lys His Glu Arg Gly Arg Pro Leu Leu Phe Lys Thr Thr Glu Gly
130 135 140

Ile Asn Lys Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Glu
145 150 155 160

Asp Thr Val Thr Tyr Lys Cys Pro Leu Leu Val Asn Thr Glu Pro Glu
165 170 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Met Tyr Gly
180 185 190

Thr Cys Thr Gln Ser Gly Glu Arg Arg Arg Glu Lys Arg Ser Val Ala
195 200 205

Leu Thr Pro His Ser Gly Met Gly Leu Glu Thr Arg Ala Glu Thr Trp
210 215 220

Met Ser Ser Glu Gly Ala Trp Lys His Ala Gln Arg Val Glu Ser Trp
225 230 235 240

Ile Leu Arg Asn Pro Gly Phe Ala Leu Leu Ala Gly Phe Met Ala Tyr
245 250 255

Met Ile Gly Gln Thr Gly Ile Gln Arg Thr Val Phe Phe Val Leu Met
260 265 270

Met Leu Val Ala Pro Ser Tyr Gly Met Arg Cys Val Gly Val Gly Asn
275 280 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ala Trp Val Asp Leu Val
290 295 300

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Gln Gly Lys Pro Thr
305 310 315 320

Leu Asp Phe Glu Leu Thr Lys Thr Thr Ala Lys Glu Val Ala Leu Leu
325 330 335

```
Arg Thr Tyr Cys Ile Glu Ala Ser Ile Ser Asn Ile Thr Thr Ala Thr
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Pro Tyr Leu Lys Glu Glu Gln Asp Gln
        355                 360                 365

Gln Tyr Ile Cys Arg Arg Asp Val Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Gly Val Val Thr Cys Ala Lys Phe Ser
385                 390                 395                 400

Cys Ser Gly Lys Ile Thr Gly Asn Leu Val Gln Ile Glu Asn Leu Glu
                405                 410                 415

Tyr Thr Val Val Val Thr Val His Asn Gly Asp Thr His Ala Val Gly
            420                 425                 430

Asn Asp Thr Ser Asn His Gly Val Thr Ala Thr Ile Thr Pro Arg Ser
        435                 440                 445

Pro Ser Val Glu Val Lys Leu Pro Asp Tyr Gly Glu Leu Thr Leu Asp
    450                 455                 460

Cys Glu Pro Arg Ser Gly Ile Asp Phe Asn Glu Met Ile Leu Met Lys
465                 470                 475                 480

Met Lys Lys Lys Thr Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
                485                 490                 495

Pro Leu Pro Trp Thr Ala Gly Ala Asp Thr Ser Glu Val His Trp Asn
            500                 505                 510

Tyr Lys Glu Arg Met Val Thr Phe Lys Val Pro His Ala Lys Arg Gln
        515                 520                 525

Asp Val Thr Val Leu Gly Ser Gln Glu Gly Ala Met His Ser Ala Leu
    530                 535                 540
```

Ala Gly Ala Thr Glu Val Asp Ser Gly Asp Gly Asn His Met Phe Ala
545 550 555 560

Gly His Leu Lys Cys Lys Val Arg Met Glu Lys Leu Arg Ile Lys Gly
565 570 575

Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys Glu Met
580 585 590

Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys Tyr Glu Gly
595 600 605

Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val Asn Lys
610 615 620

Glu Lys Val Val Gly Arg Ile Ile Ser Ser Thr Pro Leu Ala Glu Asn
625 630 635 640

Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Val Ile Gly Val Gly Asn Ser Ala Leu Thr Leu His Trp Phe
660 665 670

Arg Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ser Thr Tyr Arg Gly
675 680 685

Ala Lys Arg Met Ala Ile Leu Gly Glu Thr Ala Trp Asp Phe Gly Ser
690 695 700

Val Gly Gly Leu Phe Thr Ser Leu Gly Lys Ala Val His Gln Val Phe
705 710 715 720

Gly Ser Val Tyr Thr Thr Leu Phe Gly Gly Val Ser Trp Met Ile Arg
725 730 735

Ile Leu Ile Gly Phe Leu Val Leu Trp Ile Gly Thr Asn Ser Arg Asn
740 745 750

Thr Ser Met Ala Met Thr Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
850 855 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865 870 875 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
885 890 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
915 920 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
930 935 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945 950 955 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
965 970 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
980 985 990

Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser
995 1000 1005

Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala
1010 1015 1020

Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln
1025 1030 1035

Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp
1040 1045 1050

Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn
1055 1060 1065

Arg Gly Pro Ser Leu Arg Thr Thr Thr Ala Ser Gly Lys Leu Ile
1070 1075 1080

Thr Glu Trp Cys Cys Arg Ser Cys Thr Leu Pro Pro Leu Arg Tyr
1085 1090 1095

Arg Gly Glu Asp Gly Cys Trp Tyr Gly Met Glu Ile Arg Pro Leu
1100 1105 1110

Lys Glu Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly
1115 1120 1125

His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala
1130 1135 1140

Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His
1145 1150 1155

```
Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                 1165                 1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                 1180                 1185

Ala Thr  Met Thr Gly Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                 1195                 1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                 1210                 1215

Leu Leu  Arg Lys Leu Thr Ser  Lys Glu Leu Met Met  Thr Thr Ile
    1220                 1225                 1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Ile Pro Glu  Thr Ile Leu
    1235                 1240                 1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                 1255                 1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                 1270                 1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                 1285                 1290

Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                 1300                 1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                 1315                 1320

Thr Ile  Lys Gly Leu Asn Pro  Thr Ala Ile Phe Leu  Thr Thr Leu
    1325                 1330                 1335

Ser Arg  Thr Ser Lys Lys Arg  Ser Trp Pro Leu Asn  Glu Ala Ile
    1340                 1345                 1350
```

Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys
1355 1360 1365

Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu
1370 1375 1380

Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu
1385 1390 1395

Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser
1400 1405 1410

Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser
1415 1420 1425

Met Ser Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu
1430 1435 1440

Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
1445 1450 1455

Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys
1460 1465 1470

Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met
1475 1480 1485

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys
1490 1495 1500

Gly Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu
1505 1510 1515

Gly Thr Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu
1520 1525 1530

Met His Lys Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys
1535 1540 1545

```
Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                1555                1560

Trp Lys  Glu Gly Glu Glu Val  Gln Val Leu Ala Leu  Glu Pro Gly
    1565                1570                1575

Lys Asn  Pro Arg Ala Val Gln  Thr Lys Pro Gly Leu  Phe Lys Thr
    1580                1585                1590

Asn Ala  Gly Thr Ile Gly Ala  Val Ser Leu Asp Phe  Ser Pro Gly
    1595                1600                1605

Thr Ser  Gly Ser Pro Ile Ile  Asp Lys Lys Gly Lys  Val Val Gly
    1610                1615                1620

Leu Tyr  Gly Asn Gly Val Val  Thr Arg Ser Gly Ala  Tyr Val Ser
    1625                1630                1635

Ala Ile  Ala Gln Thr Glu Lys  Ser Ile Glu Asp Asn  Pro Glu Ile
    1640                1645                1650

Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                1660                1665

His Pro  Gly Ala Gly Lys Thr  Lys Arg Tyr Leu Pro  Ala Ile Val
    1670                1675                1680

Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                1690                1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                1705                1710

Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                1720                1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                1735                1740
```

```
Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                1750                1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                1765                1770

Tyr Ile  Ser Thr Arg Val Glu  Met Gly Glu Ala Ala  Gly Ile Phe
    1775                1780                1785

Met Thr  Ala Thr Pro Pro Gly  Ser Arg Asp Pro Phe  Pro Gln Ser
    1790                1795                1800

Asn Ala  Pro Ile Ile Asp Glu  Glu Arg Glu Ile Pro  Glu Arg Ser
    1805                1810                1815

Trp Asn  Ser Gly His Glu Trp  Val Thr Asp Phe Lys  Gly Lys Thr
    1820                1825                1830

Val Trp  Phe Val Pro Ser Ile  Lys Ala Gly Asn Asp  Ile Ala Ala
    1835                1840                1845

Cys Leu  Arg Lys Asn Gly Lys  Lys Val Ile Gln Leu  Ser Arg Lys
    1850                1855                1860

Thr Phe  Asp Ser Glu Tyr Val  Lys Thr Arg Thr Asn  Asp Trp Asp
    1865                1870                1875

Phe Val  Val Thr Thr Asp Ile  Ser Glu Met Gly Ala  Asn Phe Lys
    1880                1885                1890

Ala Glu  Arg Val Ile Asp Pro  Arg Arg Cys Met Lys  Pro Val Ile
    1895                1900                1905

Leu Thr  Asp Gly Glu Glu Arg  Val Ile Leu Ala Gly  Pro Met Pro
    1910                1915                1920

Val Thr  His Ser Ser Ala Ala  Gln Arg Arg Gly Arg  Ile Gly Arg
    1925                1930                1935
```

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
1955 1960 1965

Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met
1970 1975 1980

Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr
1985 1990 1995

Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg
2000 2005 2010

Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu
2015 2020 2025

Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys
2030 2035 2040

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met
2090 2095 2100

Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Val Arg Asp Ala Leu
2105 2110 2115

Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala
2120 2125 2130

Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu
2135 2140 2145

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

```
Ile Gln  Glu Val Asp Arg Thr  Leu Ala Lys Glu Gly  Ile Lys Arg
    2525                2530                2535

Gly Glu  Thr Asp His His Ala  Val Ser Arg Gly Ser  Ala Lys Leu
    2540                2545                2550

Arg Trp  Phe Val Glu Arg Asn  Met Val Thr Pro Glu  Gly Lys Val
    2555                2560                2565

Val Asp  Leu Gly Cys Gly Arg  Gly Gly Trp Ser Tyr  Tyr Cys Gly
    2570                2575                2580

Gly Leu  Lys Asn Val Arg Glu  Val Lys Gly Leu Thr  Lys Gly Gly
    2585                2590                2595

Pro Gly  His Glu Glu Pro Ile  Pro Met Ser Thr Tyr  Gly Trp Asn
    2600                2605                2610

Leu Val  Arg Leu Gln Ser Gly  Val Asp Val Phe Phe  Ile Pro Pro
    2615                2620                2625

Glu Lys  Cys Asp Thr Leu Leu  Cys Asp Ile Gly Glu  Ser Ser Pro
    2630                2635                2640

Asn Pro  Thr Val Glu Ala Gly  Arg Thr Leu Arg Val  Leu Asn Leu
    2645                2650                2655

Val Glu  Asn Trp Leu Asn Asn  Asn Thr Gln Phe Cys  Ile Lys Val
    2660                2665                2670

Leu Asn  Pro Tyr Met Pro Ser  Val Ile Glu Lys Met  Glu Ala Leu
    2675                2680                2685

Gln Arg  Lys Tyr Gly Gly Ala  Leu Val Arg Asn Pro  Leu Ser Arg
    2690                2695                2700

Asn Ser  Thr His Glu Met Tyr  Trp Val Ser Asn Ala  Ser Gly Asn
    2705                2710                2715
```

```
Ile Val  Ser Ser Val Asn Met  Ile Ser Arg Met Leu  Ile Asn Arg
    2720                2725                2730

Phe Thr  Met Arg Tyr Lys Lys  Ala Thr Tyr Glu Pro  Asp Val Asp
    2735                2740                2745

Leu Gly  Ser Gly Thr Arg Asn  Ile Gly Ile Glu Ser  Glu Ile Pro
    2750                2755                2760

Asn Leu  Asp Ile Ile Gly Lys  Arg Ile Glu Lys Ile  Lys Gln Glu
    2765                2770                2775

His Glu  Thr Ser Trp His Tyr  Asp Gln Asp His Pro  Tyr Lys Thr
    2780                2785                2790

Trp Ala  Tyr His Gly Ser Tyr  Glu Thr Lys Gln Thr  Gly Ser Ala
    2795                2800                2805

Ser Ser  Met Val Asn Gly Val  Val Arg Leu Leu Thr  Lys Pro Trp
    2810                2815                2820

Asp Val  Val Pro Met Val Thr  Gln Met Ala Met Thr  Asp Thr Thr
    2825                2830                2835

Pro Phe  Gly Gln Gln Arg Val  Phe Lys Glu Lys Val  Asp Thr Arg
    2840                2845                2850

Thr Gln  Glu Pro Lys Glu Gly  Thr Lys Lys Leu Met  Lys Ile Thr
    2855                2860                2865

Ala Glu  Trp Leu Trp Lys Glu  Leu Gly Lys Lys Lys  Thr Pro Arg
    2870                2875                2880

Met Cys  Thr Arg Glu Glu Phe  Thr Arg Lys Val Arg  Ser Asn Ala
    2885                2890                2895

Ala Leu  Gly Ala Ile Phe Thr  Asp Glu Asn Lys Trp  Lys Ser Ala
    2900                2905                2910
```

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys
2945 2950 2955

Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg
2960 2965 2970

Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp
2975 2980 2985

Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu
2990 2995 3000

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
3080 3085 3090

Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala
3095 3100 3105

Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile
3110 3115 3120

Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
3125 3130 3135

Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly
3140 3145 3150

Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
3155 3160 3165

Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln
3170 3175 3180

Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro
3185 3190 3195

Phe Cys Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg
3200 3205 3210

Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg
3215 3220 3225

Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala
3230 3235 3240

Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe
3245 3250 3255

His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
3260 3265 3270

Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
3275 3280 3285

His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
3290 3295 3300

Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 27
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 4, MVS

<220>
<221> dấu hiệu hỗn tạp
<222> (3773)..(3773)
<223> n là adenin hoặc guanin

<400> 27
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gacctttaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagctctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgatg gcgaacccct catgatagtg 480

gcaaaacatg aaaggggggag acctctcttg tttaagacaa cagaggggat caacaaatgc 540

actctcattg ccatggactt gggtgaaatg tgtgaggaca ctgtcacgta taaatgcccc 600

ttactggtca ataccgaacc tgaagacatt gattgctggt gcaatctcac gtctacctgg 660

gtcatgtatg ggacatgcac ccagagcgga gaacggagac gagagaagcg ctcagtagct 720

ttaacaccac attcaggaat gggattggaa acaagagctg agacatggat gtcatcggaa 780

ggggcttgga agcatgctca gagagtagag agctggatac tcagaaaccc aggattcgcg 840

ctcttggcag gatttatggc ttatatgatt gggcaaacag gaatccagcg aactgtcttc 900

tttgtcctaa tgatgctggt cgccccatcc tacggaatgc gatgcgtagg agtaggaaac 960

agagactttg tggaaggagt ctcaggtgga gcatgggtcg atctggtgct agaacatgga 1020

ggatgcgtca caaccatggc ccagggaaaa ccaaccttgg attttgaact gactaagaca 1080

acagccaagg aagtggctct gttaagaacc tattgcattg aagcctcaat atcaaacata 1140

accacggcaa caagatgtcc aacgcaagga gagccttatc taaaagagga acaagaccaa 1200

cagtacattt gccggagaga tgtggtagac agagggtggg gcaatggctg tggcttgttt 1260

ggaaaaggag gagttgtgac atgtgcgaag ttttcatgtt cggggaagat aacaggcaat 1320

ttggtccaaa ttgagaacct tgaatacaca gtggttgtaa cagtccacaa tggagacacc 1380

catgcagtag gaaatgacac gtccaatcat ggagttacag ccacgataac tcccaggtca 1440

ccatcggtgg aagtcaaatt gccggactat ggagaactaa cactcgattg tgaacccagg 1500

tctggaattg actttaatga gatgattctg atgaaaatga aaaagaaaac atggcttgtg 1560

cataagcaat ggtttttgga tctacctcta ccatggacag caggagcaga cacatcagag 1620

gttcactgga attacaaaga gagaatggtg acatttaagg ttcctcatgc caagagacag 1680

gatgtgacag tgctgggatc tcaggaagga gccatgcatt ctgccctcgc tggagccaca 1740

gaagtggact ccggtgatgg aaatcacatg tttgcaggac atctcaagtg caaagtccgt 1800

atggagaaat tgagaatcaa gggaatgtca tacacgatgt gttcaggaaa gttctcaatt 1860

gacaaagaga tggcagaaac acagcatggg acaacagtgg tgaaagtcaa gtatgaaggt 1920

gctggagctc cgtgtaaagt ccccatagag ataagagatg tgaacaagga aaaagtggtt 1980

gggcgtatca tctcatccac ccctttggct gagaatacca acagtgtaac caacatagag 2040

ttagaacccc cctttgggga cagctacata gtgataggtg ttggaaacag tgcattaaca 2100

ctccattggt tcaggaaagg gagttccatt ggcaagatgt ttgagtccac atacagaggt 2160

gcaaaacgaa tggccattct aggtgaaaca gcttgggatt ttggttccgt tggtggactg 2220

ttcacatcat tgggaaaggc tgtgcaccag gtttttggaa gtgtgtatac aaccctgttt 2280

ggaggagtct catggatgat tagaatccta attgggttcc tagtgttgtg gattggcacg 2340

aactcaagga acacttcaat ggctatgacg tgcatagctg ccggcattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cgggtgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc canggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga tccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggtaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagccatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttccc tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 28
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 4, MVS

<220>
<221> dấu hiệu hỗn tạp
<222> (1226)..(1226)
<223> Xaa là lysin hoặc arginin

<400> 28

```
Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1               5                  10                  15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
            20                  25                  30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
        35                  40                  45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
    50                  55                  60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65                  70                  75                  80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
                85                  90                  95

Arg Arg Arg Ser Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
            100                 105                 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Leu Met Ile Val
        115                 120                 125

Ala Lys His Glu Arg Gly Arg Pro Leu Leu Phe Lys Thr Thr Glu Gly
    130                 135                 140

Ile Asn Lys Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Glu
145                 150                 155                 160

Asp Thr Val Thr Tyr Lys Cys Pro Leu Leu Val Asn Thr Glu Pro Glu
                165                 170                 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Met Tyr Gly
            180                 185                 190

Thr Cys Thr Gln Ser Gly Glu Arg Arg Arg Glu Lys Arg Ser Val Ala
        195                 200                 205
```

```
Leu Thr Pro His Ser Gly Met Gly Leu Glu Thr Arg Ala Glu Thr Trp
    210                 215                 220

Met Ser Ser Glu Gly Ala Trp Lys His Ala Gln Arg Val Glu Ser Trp
225                 230                 235                 240

Ile Leu Arg Asn Pro Gly Phe Ala Leu Leu Ala Gly Phe Met Ala Tyr
                245                 250                 255

Met Ile Gly Gln Thr Gly Ile Gln Arg Thr Val Phe Phe Val Leu Met
            260                 265                 270

Met Leu Val Ala Pro Ser Tyr Gly Met Arg Cys Val Gly Val Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ala Trp Val Asp Leu Val
    290                 295                 300

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Gln Gly Lys Pro Thr
305                 310                 315                 320

Leu Asp Phe Glu Leu Thr Lys Thr Thr Ala Lys Glu Val Ala Leu Leu
                325                 330                 335

Arg Thr Tyr Cys Ile Glu Ala Ser Ile Ser Asn Ile Thr Thr Ala Thr
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Pro Tyr Leu Lys Glu Glu Gln Asp Gln
        355                 360                 365

Gln Tyr Ile Cys Arg Arg Asp Val Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Gly Val Val Thr Cys Ala Lys Phe Ser
385                 390                 395                 400

Cys Ser Gly Lys Ile Thr Gly Asn Leu Val Gln Ile Glu Asn Leu Glu
                405                 410                 415
```

Tyr Thr Val Val Val Thr Val His Asn Gly Asp Thr His Ala Val Gly
420 425 430

Asn Asp Thr Ser Asn His Gly Val Thr Ala Thr Ile Thr Pro Arg Ser
435 440 445

Pro Ser Val Glu Val Lys Leu Pro Asp Tyr Gly Glu Leu Thr Leu Asp
450 455 460

Cys Glu Pro Arg Ser Gly Ile Asp Phe Asn Glu Met Ile Leu Met Lys
465 470 475 480

Met Lys Lys Lys Thr Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
485 490 495

Pro Leu Pro Trp Thr Ala Gly Ala Asp Thr Ser Glu Val His Trp Asn
500 505 510

Tyr Lys Glu Arg Met Val Thr Phe Lys Val Pro His Ala Lys Arg Gln
515 520 525

Asp Val Thr Val Leu Gly Ser Gln Glu Gly Ala Met His Ser Ala Leu
530 535 540

Ala Gly Ala Thr Glu Val Asp Ser Gly Asp Gly Asn His Met Phe Ala
545 550 555 560

Gly His Leu Lys Cys Lys Val Arg Met Glu Lys Leu Arg Ile Lys Gly
565 570 575

Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys Glu Met
580 585 590

Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys Tyr Glu Gly
595 600 605

Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val Asn Lys
610 615 620

Glu Lys Val Val Gly Arg Ile Ile Ser Ser Thr Pro Leu Ala Glu Asn
625 630 635 640

Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Val Ile Gly Val Gly Asn Ser Ala Leu Thr Leu His Trp Phe
660 665 670

Arg Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ser Thr Tyr Arg Gly
675 680 685

Ala Lys Arg Met Ala Ile Leu Gly Glu Thr Ala Trp Asp Phe Gly Ser
690 695 700

Val Gly Gly Leu Phe Thr Ser Leu Gly Lys Ala Val His Gln Val Phe
705 710 715 720

Gly Ser Val Tyr Thr Thr Leu Phe Gly Gly Val Ser Trp Met Ile Arg
725 730 735

Ile Leu Ile Gly Phe Leu Val Leu Trp Ile Gly Thr Asn Ser Arg Asn
740 745 750

Thr Ser Met Ala Met Thr Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

```
Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
        835                 840                 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
    850                 855                 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865                 870                 875                 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
                885                 890                 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
            900                 905                 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
        915                 920                 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
            980                 985                 990

Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035
```

```
Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050

Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065

Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                1075                1080

Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                1090                1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                1105                1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                1120                1125

His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                1135                1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                1150                1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                1165                1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                1180                1185

Ala Thr  Met Thr Gly Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                1195                1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                1210                1215

Leu Leu  Arg Lys Leu Thr Ser  Xaa Glu Leu Met Met  Thr Thr Ile
    1220                1225                1230
```

```
Gly Ile  Val Leu Leu Ser Gln  Ser Thr Ile Pro Glu  Thr Ile Leu
    1235                 1240                 1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                 1255                 1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                 1270                 1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                 1285                 1290

Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                 1300                 1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                 1315                 1320

Thr Ile  Lys Gly Leu Asn Pro  Thr Ala Ile Phe Leu  Thr Thr Leu
    1325                 1330                 1335

Ser Arg  Thr Ser Lys Lys Arg  Ser Trp Pro Leu Asn  Glu Ala Ile
    1340                 1345                 1350

Met Ala  Val Gly Met Val Ser  Ile Leu Ala Ser Ser  Leu Leu Lys
    1355                 1360                 1365

Asn Asp  Ile Pro Met Thr Gly  Pro Leu Val Ala Gly  Gly Leu Leu
    1370                 1375                 1380

Thr Val  Cys Tyr Val Leu Thr  Gly Arg Ser Ala Asp  Leu Glu Leu
    1385                 1390                 1395

Glu Arg  Ala Ala Asp Val Lys  Trp Glu Asp Gln Ala  Glu Ile Ser
    1400                 1405                 1410

Gly Ser  Ser Pro Ile Leu Ser  Ile Thr Ile Ser Glu  Asp Gly Ser
    1415                 1420                 1425
```

Met Ser Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu
1430 1435 1440

Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
1445 1450 1455

Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys
1460 1465 1470

Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met
1475 1480 1485

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys
1490 1495 1500

Gly Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu
1505 1510 1515

Gly Thr Phe His Thr Met Trp His Val Thr Arg Gly Ala Val Leu
1520 1525 1530

Met His Lys Gly Lys Arg Ile Glu Pro Ser Trp Ala Asp Val Lys
1535 1540 1545

Lys Asp Leu Ile Ser Tyr Gly Gly Gly Trp Lys Leu Glu Gly Glu
1550 1555 1560

Trp Lys Glu Gly Glu Glu Val Gln Val Leu Ala Leu Glu Pro Gly
1565 1570 1575

Lys Asn Pro Arg Ala Val Gln Thr Lys Pro Gly Leu Phe Lys Thr
1580 1585 1590

Asn Ala Gly Thr Ile Gly Ala Val Ser Leu Asp Phe Ser Pro Gly
1595 1600 1605

Thr Ser Gly Ser Pro Ile Ile Asp Lys Lys Gly Lys Val Val Gly
1610 1615 1620

Leu Tyr Gly Asn Gly Val Val Thr Arg Ser Gly Ala Tyr Val Ser
1625 1630 1635

Ala Ile Ala Gln Thr Glu Lys Ser Ile Glu Asp Asn Pro Glu Ile
1640 1645 1650

Glu Asp Asp Ile Phe Arg Lys Arg Arg Leu Thr Ile Met Asp Leu
1655 1660 1665

His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val
1670 1675 1680

Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro
1685 1690 1695

Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu
1700 1705 1710

Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly
1715 1720 1725

Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg
1730 1735 1740

Leu Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met
1745 1750 1755

Asp Glu Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly
1760 1765 1770

Tyr Ile Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe
1775 1780 1785

Met Thr Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser
1790 1795 1800

Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser
1805 1810 1815

Trp Asn Ser Gly His Glu Trp Val Thr Asp Phe Lys Gly Lys Thr
1820 1825 1830

Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala
1835 1840 1845

Cys Leu Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg Lys
1850 1855 1860

Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp
1865 1870 1875

Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys
1880 1885 1890

Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile
1895 1900 1905

Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro
1910 1915 1920

Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
1925 1930 1935

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
1955 1960 1965

Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met
1970 1975 1980

Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr
1985 1990 1995

Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg
2000 2005 2010

Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu
2015 2020 2025

Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys
2030 2035 2040

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met
2090 2095 2100

Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Val Arg Asp Ala Leu
2105 2110 2115

Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala
2120 2125 2130

Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu
2135 2140 2145

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg
2525 2530 2535

Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu
2540 2545 2550

Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val
2555 2560 2565

Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly
2570 2575 2580

Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly
2585 2590 2595

Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn
2600 2605 2610

Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro
2615 2620 2625

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

```
Trp Ala  Tyr His Gly Ser Tyr  Glu Thr Lys Gln Thr  Gly Ser Ala
    2795                 2800                 2805

Ser Ser  Met Val Asn Gly Val  Val Arg Leu Leu Thr  Lys Pro Trp
    2810                 2815                 2820

Asp Val  Val Pro Met Val Thr  Gln Met Ala Met Thr  Asp Thr Thr
    2825                 2830                 2835

Pro Phe  Gly Gln Gln Arg Val  Phe Lys Glu Lys Val  Asp Thr Arg
    2840                 2845                 2850

Thr Gln  Glu Pro Lys Glu Gly  Thr Lys Lys Leu Met  Lys Ile Thr
    2855                 2860                 2865

Ala Glu  Trp Leu Trp Lys Glu  Leu Gly Lys Lys Lys  Thr Pro Arg
    2870                 2875                 2880

Met Cys  Thr Arg Glu Glu Phe  Thr Arg Lys Val Arg  Ser Asn Ala
    2885                 2890                 2895

Ala Leu  Gly Ala Ile Phe Thr  Asp Glu Asn Lys Trp  Lys Ser Ala
    2900                 2905                 2910

Arg Glu  Ala Val Glu Asp Ser  Arg Phe Trp Glu Leu  Val Asp Lys
    2915                 2920                 2925

Glu Arg  Asn Leu His Leu Glu  Gly Lys Cys Glu Thr  Cys Val Tyr
    2930                 2935                 2940

Asn Met  Met Gly Lys Arg Glu  Lys Lys Leu Gly Glu  Phe Gly Lys
    2945                 2950                 2955

Ala Lys  Gly Ser Arg Ala Ile  Trp Tyr Met Trp Leu  Gly Ala Arg
    2960                 2965                 2970

Phe Leu  Glu Phe Glu Ala Leu  Gly Phe Leu Asn Glu  Asp His Trp
    2975                 2980                 2985
```

```
Phe Ser  Arg Glu Asn Ser Leu  Ser Gly Val Glu Gly  Glu Gly Leu
    2990                 2995                 3000

His Lys  Leu Gly Tyr Ile Leu  Arg Asp Val Ser Lys  Lys Glu Gly
    3005                 3010                 3015

Gly Ala  Met Tyr Ala Asp Asp  Thr Ala Gly Trp Asp  Thr Arg Ile
    3020                 3025                 3030

Thr Leu  Glu Asp Leu Lys Asn  Glu Glu Met Val Thr  Asn His Met
    3035                 3040                 3045

Glu Gly  Glu His Lys Lys Leu  Ala Glu Ala Ile Phe  Lys Leu Thr
    3050                 3055                 3060

Tyr Gln  Asn Lys Val Val Arg  Val Gln Arg Pro Thr  Pro Arg Gly
    3065                 3070                 3075

Thr Val  Met Asp Ile Ile Ser  Arg Arg Asp Gln Arg  Gly Ser Gly
    3080                 3085                 3090

Gln Val  Gly Thr Tyr Gly Leu  Asn Thr Phe Thr Asn  Met Glu Ala
    3095                 3100                 3105

Gln Leu  Ile Arg Gln Met Glu  Gly Glu Gly Val Phe  Lys Ser Ile
    3110                 3115                 3120

Gln His  Leu Thr Ile Thr Glu  Glu Ile Ala Val Gln  Asn Trp Leu
    3125                 3130                 3135

Ala Arg  Val Gly Arg Glu Arg  Leu Ser Arg Met Ala  Ile Ser Gly
    3140                 3145                 3150

Asp Asp  Cys Val Val Lys Pro  Leu Asp Asp Arg Phe  Ala Ser Ala
    3155                 3160                 3165

Leu Thr  Ala Leu Asn Asp Met  Gly Lys Ile Arg Lys  Asp Ile Gln
    3170                 3175                 3180
```

```
Gln Trp  Glu Pro Ser Arg Gly  Trp Asn Asp Trp Thr  Gln Val Pro
    3185                 3190                 3195

Phe Cys  Ser His His Phe His  Glu Leu Ile Met Lys  Asp Gly Arg
    3200                 3205                 3210

Val Leu  Val Val Pro Cys Arg  Asn Gln Asp Glu Leu  Ile Gly Arg
    3215                 3220                 3225

Ala Arg  Ile Ser Gln Gly Ala  Gly Trp Ser Leu Arg  Glu Thr Ala
    3230                 3235                 3240

Cys Leu  Gly Lys Ser Tyr Ala  Gln Met Trp Ser Leu  Met Tyr Phe
    3245                 3250                 3255

His Arg  Arg Asp Leu Arg Leu  Ala Ala Asn Ala Ile  Cys Ser Ala
    3260                 3265                 3270

Val Pro  Ser His Trp Val Pro  Thr Ser Arg Thr Thr  Trp Ser Ile
    3275                 3280                 3285

His Ala  Lys His Glu Trp Met  Thr Thr Glu Asp Met  Leu Thr Val
    3290                 3295                 3300

Trp Asn  Arg Val Trp Ile Gln  Glu Asn Pro Trp Met  Glu Asp Lys
    3305                 3310                 3315

Thr Pro  Val Glu Ser Trp Glu  Glu Ile Pro Tyr Leu  Gly Lys Arg
    3320                 3325                 3330

Glu Asp  Gln Trp Cys Gly Ser  Leu Ile Gly Leu Thr  Ser Arg Ala
    3335                 3340                 3345

Thr Trp  Ala Lys Asn Ile Gln  Ala Ala Ile Asn Gln  Val Arg Ser
    3350                 3355                 3360

Leu Ile  Gly Asn Glu Glu Tyr  Thr Asp Tyr Met Pro  Ser Met Lys
    3365                 3370                 3375
```

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp
3380 3385 3390

<210> 29
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 4, WVS

<220>
<221> dấu hiệu hỗn tạp
<222> (3773)..(3773)
<223> n là adenin hoặc guanin

<220>
<221> dấu hiệu hỗn tạp
<222> (7538)..(7538)
<223> n là xytosin hoặc thymin

<400> 29
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gacctttaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagctctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgatg gcgaacccct catgatagtg 480

gcaaaacatg aaaggggggag acctctcttg tttaagacaa cagaggggat caacaaatgc 540

actctcattg ccatggactt gggtgaaatg tgtgaggaca ctgtcacgta taaatgcccc 600

ttactggtca ataccgaacc tgaagacatt gattgctggt gcaatctcac gtctacctgg 660

gtcatgtatg ggacatgcac ccagagcgga gaacggagac gagagaagcg ctcagtagct 720

ttaacaccac attcaggaat gggattggaa acaagagctg agacatggat gtcatcggaa 780

ggggcttgga agcatgctca gagagtagag agctggatac tcagaaaccc aggattcgcg 840

ctcttggcag gatttatggc ttatatgatt gggcaaacag gaatccagcg aactgtcttc 900

tttgtcctaa tgatgctggt cgccccatcc tacggaatgc gatgcgtagg agtaggaaac 960

agagactttg tggaaggagt ctcaggtgga gcatgggtcg atctggtgct agaacatgga 1020

ggatgcgtca caaccatggc ccagggaaaa ccaaccttgg attttgaact gactaagaca 1080

acagccaagg aagtggctct gttaagaacc tattgcattg aagcctcaat atcaaacata 1140

accacggcaa caagatgtcc aacgcaagga gagccttatc taaaagagga acaagaccaa 1200

cagtacattt gccggagaga tgtggtagac agagggtggg gcaatggctg tggcttgttt 1260

ggaaaaggag gagttgtgac atgtgcgaag ttttcatgtt cggggaagat aacaggcaat 1320

ttggtccaaa ttgagaacct tgaatacaca gtggttgtaa cagtccacaa tggagacacc 1380

catgcagtag gaaatgacac gtccaatcat ggagttacag ccacgataac tcccaggtca 1440

ccatcggtgg aagtcaaatt gccggactat ggagaactaa cactcgattg tgaacccagg 1500

tctggaattg actttaatga gatgattctg atgaaaatga aaaagaaaac atggcttgtg 1560

cataagcaat ggtttttgga tctacctcta ccatggacag caggagcaga cacatcagag 1620

gttcactgga attacaaaga gagaatggtg acatttaagg ttcctcatgc caagagacag 1680

gatgtgacag tgctgggatc tcaggaagga gccatgcatt ctgccctcgc tggagccaca 1740

gaagtggact ccggtgatgg aaatcacatg tttgcaggac atctcaagtg caaagtccgt 1800

atggagaaat tgagaatcaa gggaatgtca tacacgatgt gttcaggaaa gttctcaatt 1860

gacaaagaga tggcagaaac acagcatggg acaacagtgg tgaaagtcaa gtatgaaggt 1920

gctggagctc cgtgtaaagt ccccatagag ataagagatg tgaacaagga aaaagtggtt 1980

gggcgtatca tctcatccac ccctttggct gagaatacca acagtgtaac caacatagag 2040

ttagaacccc cctttgggga cagctacata gtgataggtg ttggaaacag tgcattaaca 2100

ctccattggt tcaggaaagg gagttccatt ggcaagatgt ttgagtccac atacagaggt 2160

gcaaaacgaa tggccattct aggtgaaaca gcttgggatt ttggttccgt tggtggactg 2220

ttcacatcat tgggaaaggc tgtgcaccag gtttttggaa gtgtgtatac aaccctgttt 2280

ggaggagtct catggatgat tagaatccta attgggttcc tagtgttgtg gattggcacg 2340

aactcaagga acacttcaat ggctatgacg tgcatagctg ccggcattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cgggtgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc canggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga tccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctccccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggtaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gtttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagccatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttnta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttccc tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 30
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 4, WVS

<220>
<221> dấu hiệu hỗn tạp
<222> (1226)..(1226)
<223> Xaa là lysin hoặc arginin

<220>
<221> dấu hiệu hỗn tạp
<222> (2481)..(2481)
<223> Xaa là serin hoặc phenylalanin

<400> 30

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

```
Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
    50                  55                  60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65                  70                  75                  80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
                85                  90                  95

Arg Arg Arg Ser Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
            100                 105                 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Leu Met Ile Val
        115                 120                 125

Ala Lys His Glu Arg Gly Arg Pro Leu Leu Phe Lys Thr Thr Glu Gly
    130                 135                 140

Ile Asn Lys Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Glu
145                 150                 155                 160

Asp Thr Val Thr Tyr Lys Cys Pro Leu Leu Val Asn Thr Glu Pro Glu
                165                 170                 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Met Tyr Gly
            180                 185                 190

Thr Cys Thr Gln Ser Gly Glu Arg Arg Arg Glu Lys Arg Ser Val Ala
        195                 200                 205

Leu Thr Pro His Ser Gly Met Gly Leu Glu Thr Arg Ala Glu Thr Trp
    210                 215                 220

Met Ser Ser Glu Gly Ala Trp Lys His Ala Gln Arg Val Glu Ser Trp
225                 230                 235                 240

Ile Leu Arg Asn Pro Gly Phe Ala Leu Leu Ala Gly Phe Met Ala Tyr
                245                 250                 255
```

Met Ile Gly Gln Thr Gly Ile Gln Arg Thr Val Phe Phe Val Leu Met
260 265 270

Met Leu Val Ala Pro Ser Tyr Gly Met Arg Cys Val Gly Val Gly Asn
275 280 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ala Trp Val Asp Leu Val
290 295 300

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Gln Gly Lys Pro Thr
305 310 315 320

Leu Asp Phe Glu Leu Thr Lys Thr Thr Ala Lys Glu Val Ala Leu Leu
325 330 335

Arg Thr Tyr Cys Ile Glu Ala Ser Ile Ser Asn Ile Thr Thr Ala Thr
340 345 350

Arg Cys Pro Thr Gln Gly Glu Pro Tyr Leu Lys Glu Glu Gln Asp Gln
355 360 365

Gln Tyr Ile Cys Arg Arg Asp Val Val Asp Arg Gly Trp Gly Asn Gly
370 375 380

Cys Gly Leu Phe Gly Lys Gly Gly Val Val Thr Cys Ala Lys Phe Ser
385 390 395 400

Cys Ser Gly Lys Ile Thr Gly Asn Leu Val Gln Ile Glu Asn Leu Glu
405 410 415

Tyr Thr Val Val Val Thr Val His Asn Gly Asp Thr His Ala Val Gly
420 425 430

Asn Asp Thr Ser Asn His Gly Val Thr Ala Thr Ile Thr Pro Arg Ser
435 440 445

Pro Ser Val Glu Val Lys Leu Pro Asp Tyr Gly Glu Leu Thr Leu Asp
450 455 460

Cys Glu Pro Arg Ser Gly Ile Asp Phe Asn Glu Met Ile Leu Met Lys
465 470 475 480

Met Lys Lys Lys Thr Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
485 490 495

Pro Leu Pro Trp Thr Ala Gly Ala Asp Thr Ser Glu Val His Trp Asn
500 505 510

Tyr Lys Glu Arg Met Val Thr Phe Lys Val Pro His Ala Lys Arg Gln
515 520 525

Asp Val Thr Val Leu Gly Ser Gln Glu Gly Ala Met His Ser Ala Leu
530 535 540

Ala Gly Ala Thr Glu Val Asp Ser Gly Asp Gly Asn His Met Phe Ala
545 550 555 560

Gly His Leu Lys Cys Lys Val Arg Met Glu Lys Leu Arg Ile Lys Gly
565 570 575

Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys Glu Met
580 585 590

Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys Tyr Glu Gly
595 600 605

Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val Asn Lys
610 615 620

Glu Lys Val Val Gly Arg Ile Ile Ser Ser Thr Pro Leu Ala Glu Asn
625 630 635 640

Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
645 650 655

Tyr Ile Val Ile Gly Val Gly Asn Ser Ala Leu Thr Leu His Trp Phe
660 665 670

```
Arg Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ser Thr Tyr Arg Gly
        675                 680                 685

Ala Lys Arg Met Ala Ile Leu Gly Glu Thr Ala Trp Asp Phe Gly Ser
    690                 695                 700

Val Gly Gly Leu Phe Thr Ser Leu Gly Lys Ala Val His Gln Val Phe
705                 710                 715                 720

Gly Ser Val Tyr Thr Thr Leu Phe Gly Gly Val Ser Trp Met Ile Arg
                725                 730                 735

Ile Leu Ile Gly Phe Leu Val Leu Trp Ile Gly Thr Asn Ser Arg Asn
            740                 745                 750

Thr Ser Met Ala Met Thr Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
        755                 760                 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
    770                 775                 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785                 790                 795                 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
                805                 810                 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
            820                 825                 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
        835                 840                 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
    850                 855                 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865                 870                 875                 880
```

```
Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
                885                 890                 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
            900                 905                 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
        915                 920                 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
            980                 985                 990

Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035

Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050

Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065

Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                1075                1080
```

```
Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                1090                1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                1105                1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                1120                1125

His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                1135                1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                1150                1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                1165                1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                1180                1185

Ala Thr  Met Thr Gly Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                1195                1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                1210                1215

Leu Leu  Arg Lys Leu Thr Ser  Xaa Glu Leu Met Met  Thr Thr Ile
    1220                1225                1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Ile Pro Glu  Thr Ile Leu
    1235                1240                1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                1255                1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                1270                1275
```

Ile Leu Cys Val Pro Asn Ala Val Ile Leu Gln Asn Ala Trp Lys
1280 1285 1290

Val Ser Cys Thr Ile Leu Ala Val Val Ser Val Ser Pro Leu Phe
1295 1300 1305

Leu Thr Ser Ser Gln Gln Lys Thr Asp Trp Ile Pro Leu Ala Leu
1310 1315 1320

Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu
1325 1330 1335

Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile
1340 1345 1350

Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys
1355 1360 1365

Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu
1370 1375 1380

Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu
1385 1390 1395

Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser
1400 1405 1410

Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser
1415 1420 1425

Met Ser Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu
1430 1435 1440

Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
1445 1450 1455

Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys
1460 1465 1470

```
Gln Arg  Ala Gly Val Leu Trp  Asp Val Pro Ser Pro  Pro Pro Met
    1475                 1480                 1485

Gly Lys  Ala Glu Leu Glu Asp  Gly Ala Tyr Arg Ile  Lys Gln Lys
    1490                 1495                 1500

Gly Ile  Leu Gly Tyr Ser Gln  Ile Gly Ala Gly Val  Tyr Lys Glu
    1505                 1510                 1515

Gly Thr  Phe His Thr Met Trp  His Val Thr Arg Gly  Ala Val Leu
    1520                 1525                 1530

Met His  Lys Gly Lys Arg Ile  Glu Pro Ser Trp Ala  Asp Val Lys
    1535                 1540                 1545

Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                 1555                 1560

Trp Lys  Glu Gly Glu Glu Val  Gln Val Leu Ala Leu  Glu Pro Gly
    1565                 1570                 1575

Lys Asn  Pro Arg Ala Val Gln  Thr Lys Pro Gly Leu  Phe Lys Thr
    1580                 1585                 1590

Asn Ala  Gly Thr Ile Gly Ala  Val Ser Leu Asp Phe  Ser Pro Gly
    1595                 1600                 1605

Thr Ser  Gly Ser Pro Ile Ile  Asp Lys Lys Gly Lys  Val Val Gly
    1610                 1615                 1620

Leu Tyr  Gly Asn Gly Val Val  Thr Arg Ser Gly Ala  Tyr Val Ser
    1625                 1630                 1635

Ala Ile  Ala Gln Thr Glu Lys  Ser Ile Glu Asp Asn  Pro Glu Ile
    1640                 1645                 1650

Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                 1660                 1665
```

His Pro Gly Ala Gly Lys Thr Lys Arg Tyr Leu Pro Ala Ile Val
1670 1675 1680

Arg Glu Ala Ile Lys Arg Gly Leu Arg Thr Leu Ile Leu Ala Pro
1685 1690 1695

Thr Arg Val Val Ala Ala Glu Met Glu Glu Ala Leu Arg Gly Leu
1700 1705 1710

Pro Ile Arg Tyr Gln Thr Pro Ala Ile Arg Ala Val His Thr Gly
1715 1720 1725

Arg Glu Ile Val Asp Leu Met Cys His Ala Thr Phe Thr Met Arg
1730 1735 1740

Leu Leu Ser Pro Val Arg Val Pro Asn Tyr Asn Leu Ile Ile Met
1745 1750 1755

Asp Glu Ala His Phe Thr Asp Pro Ala Ser Ile Ala Ala Arg Gly
1760 1765 1770

Tyr Ile Ser Thr Arg Val Glu Met Gly Glu Ala Ala Gly Ile Phe
1775 1780 1785

Met Thr Ala Thr Pro Pro Gly Ser Arg Asp Pro Phe Pro Gln Ser
1790 1795 1800

Asn Ala Pro Ile Ile Asp Glu Glu Arg Glu Ile Pro Glu Arg Ser
1805 1810 1815

Trp Asn Ser Gly His Glu Trp Val Thr Asp Phe Lys Gly Lys Thr
1820 1825 1830

Val Trp Phe Val Pro Ser Ile Lys Ala Gly Asn Asp Ile Ala Ala
1835 1840 1845

Cys Leu Arg Lys Asn Gly Lys Lys Val Ile Gln Leu Ser Arg Lys
1850 1855 1860

Thr Phe Asp Ser Glu Tyr Val Lys Thr Arg Thr Asn Asp Trp Asp
1865 1870 1875

Phe Val Val Thr Thr Asp Ile Ser Glu Met Gly Ala Asn Phe Lys
1880 1885 1890

Ala Glu Arg Val Ile Asp Pro Arg Arg Cys Met Lys Pro Val Ile
1895 1900 1905

Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro
1910 1915 1920

Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
1925 1930 1935

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
1955 1960 1965

Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met
1970 1975 1980

Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr
1985 1990 1995

Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg
2000 2005 2010

Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu
2015 2020 2025

Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys
2030 2035 2040

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met
2090 2095 2100

Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Val Arg Asp Ala Leu
2105 2110 2115

Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala
2120 2125 2130

Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu
2135 2140 2145

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Thr Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Xaa Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg
2525 2530 2535

Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu
2540 2545 2550

Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val
2555 2560 2565

Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly
2570 2575 2580

Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly
2585 2590 2595

Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn
2600 2605 2610

Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro
2615 2620 2625

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp
2810 2815 2820

Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr
2825 2830 2835

Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
2840 2845 2850

Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr
2855 2860 2865

Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg
2870 2875 2880

Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
2885 2890 2895

Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala
2900 2905 2910

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys
2945 2950 2955

Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg
2960 2965 2970

Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp
2975 2980 2985

Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu
2990 2995 3000

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
3080 3085 3090

Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala
3095 3100 3105

Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile
3110 3115 3120

Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
3125 3130 3135

Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly
3140 3145 3150

Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
3155 3160 3165

Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln
3170 3175 3180

Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro
3185 3190 3195

Phe Cys Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg
3200 3205 3210

Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg
3215 3220 3225

```
Ala Arg  Ile Ser Gln Gly Ala  Gly Trp Ser Leu Arg  Glu Thr Ala
    3230                 3235                 3240

Cys Leu  Gly Lys Ser Tyr Ala  Gln Met Trp Ser Leu  Met Tyr Phe
    3245                 3250                 3255

His Arg  Arg Asp Leu Arg Leu  Ala Ala Asn Ala Ile  Cys Ser Ala
    3260                 3265                 3270

Val Pro  Ser His Trp Val Pro  Thr Ser Arg Thr Thr  Trp Ser Ile
    3275                 3280                 3285

His Ala  Lys His Glu Trp Met  Thr Thr Glu Asp Met  Leu Thr Val
    3290                 3295                 3300

Trp Asn  Arg Val Trp Ile Gln  Glu Asn Pro Trp Met  Glu Asp Lys
    3305                 3310                 3315

Thr Pro  Val Glu Ser Trp Glu  Glu Ile Pro Tyr Leu  Gly Lys Arg
    3320                 3325                 3330

Glu Asp  Gln Trp Cys Gly Ser  Leu Ile Gly Leu Thr  Ser Arg Ala
    3335                 3340                 3345

Thr Trp  Ala Lys Asn Ile Gln  Ala Ala Ile Asn Gln  Val Arg Ser
    3350                 3355                 3360

Leu Ile  Gly Asn Glu Glu Tyr  Thr Asp Tyr Met Pro  Ser Met Lys
    3365                 3370                 3375

Arg Phe  Arg Arg Glu Glu Glu  Glu Ala Gly Val Leu  Trp
    3380                 3385                 3390
```

<210> 31
<211> 10723
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 4, BVS

<220>
<221> dấu hiệu hỗn tạp
<222> (3773)..(3773)
<223> n là adenin hoặc guanin

<220>
<221> dấu hiệu hỗn tạp
<222> (7538)..(7538)
<223> n là xytosin hoặc thymin

<400> 31
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 60

gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 120

aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag 180

ctgacaaaga gattctcact tggaatgctg cagggacgag gacctttaaa actgttcatg 240

gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga 300

tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360

ggaaggatgc tgaacatctt gaataggaga cgcagctctg caggcatgat cattatgctg 420

attccaacag tgatggcgtt ccatttaacc acgcgtgatg gcgaacccct catgatagtg 480

gcaaaacatg aaaggggggag acctctcttg tttaagacaa cagaggggat caacaaatgc 540

actctcattg ccatggactt gggtgaaatg tgtgaggaca ctgtcacgta taaatgcccc 600

ttactggtca ataccgaacc tgaagacatt gattgctggt gcaatctcac gtctacctgg 660

gtcatgtatg ggacatgcac ccagagcgga gaacggagac gagagaagcg ctcagtagct 720

ttaacaccac attcaggaat gggattggaa acaagagctg agacatggat gtcatcggaa 780

ggggcttgga agcatgctca gagagtagag agctggatac tcagaaaccc aggattcgcg 840

ctcttggcag gatttatggc ttatatgatt gggcaaacag gaatccagcg aactgtcttc 900

tttgtcctaa tgatgctggt cgccccatcc tacggaatgc gatgcgtagg agtaggaaac 960

agagactttg tggaaggagt ctcaggtgga gcatgggtcg atctggtgct agaacatgga 1020

ggatgcgtca caaccatggc ccagggaaaa ccaaccttgg attttgaact gactaagaca 1080

acagccaagg aagtggctct gttaagaacc tattgcattg aagcctcaat atcaaacata 1140

accacggcaa caagatgtcc aacgcaagga gagccttatc taaaagagga acaagaccaa 1200

cagtacattt gccggagaga tgtggtagac agagggtggg gcaatggctg tggcttgttt 1260

ggaaaaggag gagttgtgac atgtgcgaag ttttcatgtt cggggaagat aacaggcaat 1320

ttggtccaaa ttgagaacct tgaatacaca gtggttgtaa cagtccacaa tggagacacc 1380

catgcagtag gaaatgacac gtccaatcat ggagttacag ccacgataac tcccaggtca 1440

ccatcggtgg aagtcaaatt gccggactat ggagaactaa cactcgattg tgaacccagg 1500

tctggaattg actttaatga gatgattctg atgaaaatga aaaagaaaac atggcttgtg 1560

cataagcaat ggtttttgga tctacctcta ccatggacag caggagcaga cacatcagag 1620

gttcactgga attacaaaga gagaatggtg acatttaagg ttcctcatgc caagagacag 1680

gatgtgacag tgctgggatc tcaggaagga gccatgcatt ctgccctcgc tggagccaca 1740

gaagtggact ccggtgatgg aaatcacatg tttgcaggac atctcaagtg caaagtccgt 1800

atggagaaat tgagaatcaa gggaatgtca tacacgatgt gttcaggaaa gttctcaatt 1860

gacaaagaga tggcagaaac acagcatggg acaacagtgg tgaaagtcaa gtatgaaggt 1920

gctggagctc cgtgtaaagt ccccatagag ataagagatg tgaacaagga aaaagtggtt 1980

gggcgtatca tctcatccac ccctttggct gagaatacca acagtgtaac caacatagag 2040

ttagaacccc cctttgggga cagctacata gtgataggtg ttggaaacag tgcattaaca 2100

ctccattggt tcaggaaagg gagttccatt ggcaagatgt ttgagtccac atacagaggt 2160

gcaaaacgaa tggccattct aggtgaaaca gcttgggatt ttggttccgt tggtggactg 2220

ttcacatcat tgggaaaggc tgtgcaccag gtttttggaa gtgtgtatac aaccctgttt 2280

ggaggagtct catggatgat tagaatccta attgggttcc tagtgttgtg gattggcacg 2340

aactcaagga acacttcaat ggctatgacg tgcatagctg ccggcattgt gacactgtat 2400

ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg 2460

aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag 2520

ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac 2580

atttgtggaa tccgctcagt aacaagactg gagaatctga tgtggaaaca aataacacca 2640

gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc 2700

aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat 2760

tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt 2820

ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg 2880

gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa 2940

aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc 3000

gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag 3060

aaagcctctt tcattgaagt taaaaactgc cactggccaa aatcacacac cctctggagc 3120

aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa 3180

cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt 3240

gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat 3300

agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc 3360

cgatcttgca cattaccacc gctaagatac agaggtgagg atgggtgctg gtacgggatg 3420

gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga 3480

catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa 3540

atgcttagga cccgagtagg aacgaaacat gcaatactac tagttgcagt ttcttttgtg 3600

acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc 3660

gccactatga cgggtgacat aggtatgggc gtgacttatc ttgccctact agcagccttc 3720

aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc canggaattg 3780

atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt 3840

gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa 3900

aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta 3960

caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc 4020

ttaacatcct cacagcaaaa aacagattgg ataccattag cattgacgat caaaggtctc 4080

aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca 4140

ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa 4200

aatgatattc ccatgacagg accattagtg gctggagggc tcctcactgt gtgctacgtg 4260

ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac 4320

caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc 4380

atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg 4440

ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg 4500

tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg 4560

ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat 4620

tcccagatcg gagccggagt ttacaaagaa ggaacattcc atacaatgtg gcatgtcaca 4680

cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag 4740

aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa 4800

gaagtccagg tattggcact ggagcctgga aaaaatccaa gagccgtcca aacgaaacct 4860

ggtcttttca aaaccaacgc cggaacaata ggtgctgtat ctctggactt ttctcctgga 4920

acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt 4980

gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa 5040

gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc 5100

cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa 5160

cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa 5220

gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg 5280

cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt 5340

agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga tccagcaagt 5400

atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt 5460

atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata 5520

gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat 5580

tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct 5640

tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag 5700

tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg 5760

ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata 5820

ctaacagatg gtgaagagcg ggtgattctg gcaggaccta tgccagtgac ccactctagt 5880

gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata 5940

tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg 6000

ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt 6060

gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt 6120

gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa 6180

ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta 6240

gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc 6300

agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt 6360

gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc 6420

ttcatgactc agaaggtaag agacgcactg gacaacttag cagtgctgca cacggctgag 6480

gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg 6540

cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca 6600

aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta 6660

tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gtttttttctc 6720

atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc 6780

tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc 6840

ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc 6900

aacatcctgg acatagatct acgtcctgca tcagcatgga cgctgtatgc cgtggccaca 6960

acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta 7020

acagccatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca 7080

aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata 7140

actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc 7200

caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca 7260

actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa 7320

aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg 7380

actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg 7440

gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt 7500

agagggagtt acttggccgg agctggactt ctcttttnta ttatgaagaa cacaaccaac 7560

acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg 7620

aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat 7680

agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga 7740

ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta 7800

gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta 7860

agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca 7920

acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca 7980

gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa 8040

gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa 8100

ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta 8160

caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag 8220

atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg 8280

atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac 8340

ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt 8400

gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac 8460

cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca 8520

tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg 8580

gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag 8640

aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca 8700

gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa 8760

gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac 8820

aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag 8880

gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa 8940

agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg 9000

tggcttggag cacgcttctt agagtttgaa gccctaggat tcttaaatga agatcactgg 9060

ttctccagag agaactccct gagtggagtg gaaggagaag ggctgcacaa gctaggttac 9120

attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga 9180

tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg 9240

gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg 9300

gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac 9360

caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc 9420

caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc 9480

acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga 9540

atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct 9600

ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca 9660

agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc 9720

atgaaagacg gtcgcgtact cgttgttccc tgtagaaacc aagatgaact gattggcaga 9780

gcccgaatct cccaaggagc agggtggtct ttgcgggaga cggcctgttt ggggaagtct 9840

tacgcccaaa tgtggagctt gatgtacttc cacagacgcg acctcaggct ggcggcaaat 9900

gctatttgct cggcagtacc atcacattgg gttccaacaa gtcgaacaac ctggtccata 9960

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020

attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080

tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140

acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200

gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260

gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320

catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380

ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440

tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500

ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560

agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620

catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680

gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

<210> 32
<211> 3391
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 4, BVS

<220>
<221> dấu hiệu hỗn tạp
<222> (1226)..(1226)
<223> Xaa là lysin hoặc arginin

<220>
<221> dấu hiệu hỗn tạp
<222> (2481)..(2481)
<223> Xaa là serin hoặc phenylalanin

<400> 32

Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1 5 10 15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
20 25 30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
35 40 45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
50 55 60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65 70 75 80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
85 90 95

```
Arg Arg Arg Ser Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
            100                 105                 110

Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Leu Met Ile Val
        115                 120                 125

Ala Lys His Glu Arg Gly Arg Pro Leu Leu Phe Lys Thr Thr Glu Gly
    130                 135                 140

Ile Asn Lys Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Glu
145                 150                 155                 160

Asp Thr Val Thr Tyr Lys Cys Pro Leu Leu Val Asn Thr Glu Pro Glu
                165                 170                 175

Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Met Tyr Gly
            180                 185                 190

Thr Cys Thr Gln Ser Gly Glu Arg Arg Arg Glu Lys Arg Ser Val Ala
        195                 200                 205

Leu Thr Pro His Ser Gly Met Gly Leu Glu Thr Arg Ala Glu Thr Trp
    210                 215                 220

Met Ser Ser Glu Gly Ala Trp Lys His Ala Gln Arg Val Glu Ser Trp
225                 230                 235                 240

Ile Leu Arg Asn Pro Gly Phe Ala Leu Leu Ala Gly Phe Met Ala Tyr
                245                 250                 255

Met Ile Gly Gln Thr Gly Ile Gln Arg Thr Val Phe Phe Val Leu Met
            260                 265                 270

Met Leu Val Ala Pro Ser Tyr Gly Met Arg Cys Val Gly Val Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ala Trp Val Asp Leu Val
    290                 295                 300
```

```
Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Gln Gly Lys Pro Thr
305                 310                 315                 320

Leu Asp Phe Glu Leu Thr Lys Thr Thr Ala Lys Glu Val Ala Leu Leu
                325                 330                 335

Arg Thr Tyr Cys Ile Glu Ala Ser Ile Ser Asn Ile Thr Thr Ala Thr
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Pro Tyr Leu Lys Glu Glu Gln Asp Gln
        355                 360                 365

Gln Tyr Ile Cys Arg Arg Asp Val Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Gly Val Val Thr Cys Ala Lys Phe Ser
385                 390                 395                 400

Cys Ser Gly Lys Ile Thr Gly Asn Leu Val Gln Ile Glu Asn Leu Glu
                405                 410                 415

Tyr Thr Val Val Val Thr Val His Asn Gly Asp Thr His Ala Val Gly
            420                 425                 430

Asn Asp Thr Ser Asn His Gly Val Thr Ala Thr Ile Thr Pro Arg Ser
        435                 440                 445

Pro Ser Val Glu Val Lys Leu Pro Asp Tyr Gly Glu Leu Thr Leu Asp
    450                 455                 460

Cys Glu Pro Arg Ser Gly Ile Asp Phe Asn Glu Met Ile Leu Met Lys
465                 470                 475                 480

Met Lys Lys Lys Thr Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
                485                 490                 495

Pro Leu Pro Trp Thr Ala Gly Ala Asp Thr Ser Glu Val His Trp Asn
            500                 505                 510
```

```
Tyr Lys Glu Arg Met Val Thr Phe Lys Val Pro His Ala Lys Arg Gln
        515                 520                 525

Asp Val Thr Val Leu Gly Ser Gln Glu Gly Ala Met His Ser Ala Leu
    530                 535                 540

Ala Gly Ala Thr Glu Val Asp Ser Gly Asp Gly Asn His Met Phe Ala
545                 550                 555                 560

Gly His Leu Lys Cys Lys Val Arg Met Glu Lys Leu Arg Ile Lys Gly
                565                 570                 575

Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys Glu Met
            580                 585                 590

Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys Tyr Glu Gly
        595                 600                 605

Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val Asn Lys
    610                 615                 620

Glu Lys Val Val Gly Arg Ile Ile Ser Ser Thr Pro Leu Ala Glu Asn
625                 630                 635                 640

Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
                645                 650                 655

Tyr Ile Val Ile Gly Val Gly Asn Ser Ala Leu Thr Leu His Trp Phe
            660                 665                 670

Arg Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ser Thr Tyr Arg Gly
        675                 680                 685

Ala Lys Arg Met Ala Ile Leu Gly Glu Thr Ala Trp Asp Phe Gly Ser
    690                 695                 700

Val Gly Gly Leu Phe Thr Ser Leu Gly Lys Ala Val His Gln Val Phe
705                 710                 715                 720
```

Gly Ser Val Tyr Thr Thr Leu Phe Gly Gly Val Ser Trp Met Ile Arg
725 730 735

Ile Leu Ile Gly Phe Leu Val Leu Trp Ile Gly Thr Asn Ser Arg Asn
740 745 750

Thr Ser Met Ala Met Thr Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr
755 760 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
770 775 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785 790 795 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
805 810 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
820 825 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
835 840 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
850 855 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865 870 875 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
885 890 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
900 905 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
915 920 925

```
Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
            980                 985                 990

Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035

Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050

Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065

Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                1075                1080

Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                1090                1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                1105                1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                1120                1125
```

```
His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                1135                1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                1150                1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                1165                1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                1180                1185

Ala Thr  Met Thr Gly Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                1195                1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                1210                1215

Leu Leu  Arg Lys Leu Thr Ser  Xaa Glu Leu Met Met  Thr Thr Ile
    1220                1225                1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Ile Pro Glu  Thr Ile Leu
    1235                1240                1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                1255                1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                1270                1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                1285                1290

Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                1300                1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                1315                1320
```

Thr Ile Lys Gly Leu Asn Pro Thr Ala Ile Phe Leu Thr Thr Leu
1325 1330 1335

Ser Arg Thr Ser Lys Lys Arg Ser Trp Pro Leu Asn Glu Ala Ile
1340 1345 1350

Met Ala Val Gly Met Val Ser Ile Leu Ala Ser Ser Leu Leu Lys
1355 1360 1365

Asn Asp Ile Pro Met Thr Gly Pro Leu Val Ala Gly Gly Leu Leu
1370 1375 1380

Thr Val Cys Tyr Val Leu Thr Gly Arg Ser Ala Asp Leu Glu Leu
1385 1390 1395

Glu Arg Ala Ala Asp Val Lys Trp Glu Asp Gln Ala Glu Ile Ser
1400 1405 1410

Gly Ser Ser Pro Ile Leu Ser Ile Thr Ile Ser Glu Asp Gly Ser
1415 1420 1425

Met Ser Ile Lys Asn Glu Glu Glu Glu Gln Thr Leu Thr Ile Leu
1430 1435 1440

Ile Arg Thr Gly Leu Leu Val Ile Ser Gly Leu Phe Pro Val Ser
1445 1450 1455

Ile Pro Ile Thr Ala Ala Ala Trp Tyr Leu Trp Glu Val Lys Lys
1460 1465 1470

Gln Arg Ala Gly Val Leu Trp Asp Val Pro Ser Pro Pro Pro Met
1475 1480 1485

Gly Lys Ala Glu Leu Glu Asp Gly Ala Tyr Arg Ile Lys Gln Lys
1490 1495 1500

Gly Ile Leu Gly Tyr Ser Gln Ile Gly Ala Gly Val Tyr Lys Glu
1505 1510 1515

```
Gly Thr  Phe His Thr Met Trp  His Val Thr Arg Gly  Ala Val Leu
    1520                1525                1530

Met His  Lys Gly Lys Arg Ile  Glu Pro Ser Trp Ala  Asp Val Lys
    1535                1540                1545

Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                1555                1560

Trp Lys  Glu Gly Glu Glu Val  Gln Val Leu Ala Leu  Glu Pro Gly
    1565                1570                1575

Lys Asn  Pro Arg Ala Val Gln  Thr Lys Pro Gly Leu  Phe Lys Thr
    1580                1585                1590

Asn Ala  Gly Thr Ile Gly Ala  Val Ser Leu Asp Phe  Ser Pro Gly
    1595                1600                1605

Thr Ser  Gly Ser Pro Ile Ile  Asp Lys Lys Gly Lys  Val Val Gly
    1610                1615                1620

Leu Tyr  Gly Asn Gly Val Val  Thr Arg Ser Gly Ala  Tyr Val Ser
    1625                1630                1635

Ala Ile  Ala Gln Thr Glu Lys  Ser Ile Glu Asp Asn  Pro Glu Ile
    1640                1645                1650

Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                1660                1665

His Pro  Gly Ala Gly Lys Thr  Lys Arg Tyr Leu Pro  Ala Ile Val
    1670                1675                1680

Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                1690                1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                1705                1710
```

```
Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                1720                1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                1735                1740

Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                1750                1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                1765                1770

Tyr Ile  Ser Thr Arg Val Glu  Met Gly Glu Ala Ala  Gly Ile Phe
    1775                1780                1785

Met Thr  Ala Thr Pro Pro Gly  Ser Arg Asp Pro Phe  Pro Gln Ser
    1790                1795                1800

Asn Ala  Pro Ile Ile Asp Glu  Glu Arg Glu Ile Pro  Glu Arg Ser
    1805                1810                1815

Trp Asn  Ser Gly His Glu Trp  Val Thr Asp Phe Lys  Gly Lys Thr
    1820                1825                1830

Val Trp  Phe Val Pro Ser Ile  Lys Ala Gly Asn Asp  Ile Ala Ala
    1835                1840                1845

Cys Leu  Arg Lys Asn Gly Lys  Lys Val Ile Gln Leu  Ser Arg Lys
    1850                1855                1860

Thr Phe  Asp Ser Glu Tyr Val  Lys Thr Arg Thr Asn  Asp Trp Asp
    1865                1870                1875

Phe Val  Val Thr Thr Asp Ile  Ser Glu Met Gly Ala  Asn Phe Lys
    1880                1885                1890

Ala Glu  Arg Val Ile Asp Pro  Arg Arg Cys Met Lys  Pro Val Ile
    1895                1900                1905
```

```
Leu Thr  Asp Gly Glu Glu Arg  Val Ile Leu Ala Gly  Pro Met Pro
    1910                 1915                 1920

Val Thr  His Ser Ser Ala Ala  Gln Arg Arg Gly Arg  Ile Gly Arg
    1925                 1930                 1935

Asn Pro  Lys Asn Glu Asn Asp  Gln Tyr Ile Tyr Met  Gly Glu Pro
    1940                 1945                 1950

Leu Glu  Asn Asp Glu Asp Cys  Ala His Trp Lys Glu  Ala Lys Met
    1955                 1960                 1965

Leu Leu  Asp Asn Ile Asn Thr  Pro Glu Gly Ile Ile  Pro Ser Met
    1970                 1975                 1980

Phe Glu  Pro Glu Arg Glu Lys  Val Asp Ala Ile Asp  Gly Glu Tyr
    1985                 1990                 1995

Arg Leu  Arg Gly Glu Ala Arg  Lys Thr Phe Val Asp  Leu Met Arg
    2000                 2005                 2010

Arg Gly  Asp Leu Pro Val Trp  Leu Ala Tyr Arg Val  Ala Ala Glu
    2015                 2020                 2025

Gly Ile  Asn Tyr Ala Asp Arg  Arg Trp Cys Phe Asp  Gly Val Lys
    2030                 2035                 2040

Asn Asn  Gln Ile Leu Glu Glu  Asn Val Glu Val Glu  Ile Trp Thr
    2045                 2050                 2055

Lys Glu  Gly Glu Arg Lys Lys  Leu Lys Pro Arg Trp  Leu Asp Ala
    2060                 2065                 2070

Arg Ile  Tyr Ser Asp Pro Leu  Ala Leu Lys Glu Phe  Lys Glu Phe
    2075                 2080                 2085

Ala Ala  Gly Arg Lys Ser Leu  Thr Leu Asn Leu Ile  Thr Glu Met
    2090                 2095                 2100
```

```
Gly Arg  Leu Pro Thr Phe Met  Thr Gln Lys Val Arg  Asp Ala Leu
    2105                2110                2115

Asp Asn  Leu Ala Val Leu His  Thr Ala Glu Ala Gly  Gly Arg Ala
    2120                2125                2130

Tyr Asn  His Ala Leu Ser Glu  Leu Pro Glu Thr Leu  Glu Thr Leu
    2135                2140                2145

Leu Leu  Leu Thr Leu Leu Ala  Thr Val Thr Gly Gly  Ile Phe Leu
    2150                2155                2160

Phe Leu  Met Ser Ala Arg Gly  Ile Gly Lys Met Thr  Leu Gly Met
    2165                2170                2175

Cys Cys  Ile Ile Thr Ala Ser  Ile Leu Leu Trp Tyr  Ala Gln Ile
    2180                2185                2190

Gln Pro  His Trp Ile Ala Ala  Ser Ile Ile Leu Glu  Phe Phe Leu
    2195                2200                2205

Ile Val  Leu Leu Ile Pro Glu  Pro Glu Lys Gln Arg  Thr Pro Gln
    2210                2215                2220

Asp Asn  Gln Leu Thr Tyr Val  Val Ile Ala Ile Leu  Thr Val Val
    2225                2230                2235

Ala Ala  Thr Met Ala Asn Glu  Met Gly Phe Leu Glu  Lys Thr Lys
    2240                2245                2250

Lys Asp  Leu Gly Leu Gly Ser  Ile Ala Thr Gln Gln  Pro Glu Ser
    2255                2260                2265

Asn Ile  Leu Asp Ile Asp Leu  Arg Pro Ala Ser Ala  Trp Thr Leu
    2270                2275                2280

Tyr Ala  Val Ala Thr Thr Phe  Val Thr Pro Met Leu  Arg His Ser
    2285                2290                2295
```

```
Ile Glu  Asn Ser Ser Val Asn  Val Ser Leu Thr Ala  Ile Ala Asn
    2300                 2305                 2310

Gln Ala  Thr Val Leu Met Gly  Leu Gly Lys Gly Trp  Pro Leu Ser
    2315                 2320                 2325

Lys Met  Asp Ile Gly Val Pro  Leu Leu Ala Ile Gly  Cys Tyr Ser
    2330                 2335                 2340

Gln Val  Asn Pro Ile Thr Leu  Thr Ala Ala Leu Phe  Leu Leu Val
    2345                 2350                 2355

Ala His  Tyr Ala Ile Ile Gly  Pro Gly Leu Gln Ala  Lys Ala Thr
    2360                 2365                 2370

Arg Glu  Ala Gln Lys Arg Ala  Ala Ala Gly Ile Met  Lys Asn Pro
    2375                 2380                 2385

Thr Val  Asp Gly Ile Thr Val  Ile Asp Leu Asp Pro  Ile Pro Tyr
    2390                 2395                 2400

Asp Pro  Lys Phe Glu Lys Gln  Leu Gly Gln Val Met  Leu Leu Val
    2405                 2410                 2415

Leu Cys  Val Thr Gln Val Leu  Met Met Arg Thr Thr  Trp Ala Leu
    2420                 2425                 2430

Cys Glu  Ala Leu Thr Leu Ala  Thr Gly Pro Ile Ser  Thr Leu Trp
    2435                 2440                 2445

Glu Gly  Asn Pro Gly Arg Phe  Trp Asn Thr Thr Ile  Ala Val Ser
    2450                 2455                 2460

Met Ala  Asn Ile Phe Arg Gly  Ser Tyr Leu Ala Gly  Ala Gly Leu
    2465                 2470                 2475

Leu Phe  Xaa Ile Met Lys Asn  Thr Thr Asn Thr Arg  Arg Gly Thr
    2480                 2485                 2490
```

```
Gly Asn  Ile Gly Glu Thr Leu  Gly Glu Lys Trp Lys  Ser Arg Leu
    2495                 2500                 2505

Asn Ala  Leu Gly Lys Ser Glu  Phe Gln Ile Tyr Lys  Lys Ser Gly
    2510                 2515                 2520

Ile Gln  Glu Val Asp Arg Thr  Leu Ala Lys Glu Gly  Ile Lys Arg
    2525                 2530                 2535

Gly Glu  Thr Asp His His Ala  Val Ser Arg Gly Ser  Ala Lys Leu
    2540                 2545                 2550

Arg Trp  Phe Val Glu Arg Asn  Met Val Thr Pro Glu  Gly Lys Val
    2555                 2560                 2565

Val Asp  Leu Gly Cys Gly Arg  Gly Gly Trp Ser Tyr  Tyr Cys Gly
    2570                 2575                 2580

Gly Leu  Lys Asn Val Arg Glu  Val Lys Gly Leu Thr  Lys Gly Gly
    2585                 2590                 2595

Pro Gly  His Glu Glu Pro Ile  Pro Met Ser Thr Tyr  Gly Trp Asn
    2600                 2605                 2610

Leu Val  Arg Leu Gln Ser Gly  Val Asp Val Phe Phe  Ile Pro Pro
    2615                 2620                 2625

Glu Lys  Cys Asp Thr Leu Leu  Cys Asp Ile Gly Glu  Ser Ser Pro
    2630                 2635                 2640

Asn Pro  Thr Val Glu Ala Gly  Arg Thr Leu Arg Val  Leu Asn Leu
    2645                 2650                 2655

Val Glu  Asn Trp Leu Asn Asn  Asn Thr Gln Phe Cys  Ile Lys Val
    2660                 2665                 2670

Leu Asn  Pro Tyr Met Pro Ser  Val Ile Glu Lys Met  Glu Ala Leu
    2675                 2680                 2685
```

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp
2810 2815 2820

Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr
2825 2830 2835

Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
2840 2845 2850

Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr
2855 2860 2865

Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg
2870 2875 2880

Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
2885 2890 2895

Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala
2900 2905 2910

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys
2945 2950 2955

Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg
2960 2965 2970

Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp
2975 2980 2985

Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu
2990 2995 3000

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

```
Thr Val  Met Asp Ile Ile Ser  Arg Arg Asp Gln Arg  Gly Ser Gly
    3080                3085                3090

Gln Val  Gly Thr Tyr Gly Leu  Asn Thr Phe Thr Asn  Met Glu Ala
    3095                3100                3105

Gln Leu  Ile Arg Gln Met Glu  Gly Glu Gly Val Phe  Lys Ser Ile
    3110                3115                3120

Gln His  Leu Thr Ile Thr Glu  Glu Ile Ala Val Gln  Asn Trp Leu
    3125                3130                3135

Ala Arg  Val Gly Arg Glu Arg  Leu Ser Arg Met Ala  Ile Ser Gly
    3140                3145                3150

Asp Asp  Cys Val Val Lys Pro  Leu Asp Asp Arg Phe  Ala Ser Ala
    3155                3160                3165

Leu Thr  Ala Leu Asn Asp Met  Gly Lys Ile Arg Lys  Asp Ile Gln
    3170                3175                3180

Gln Trp  Glu Pro Ser Arg Gly  Trp Asn Asp Trp Thr  Gln Val Pro
    3185                3190                3195

Phe Cys  Ser His His Phe His  Glu Leu Ile Met Lys  Asp Gly Arg
    3200                3205                3210

Val Leu  Val Val Pro Cys Arg  Asn Gln Asp Glu Leu  Ile Gly Arg
    3215                3220                3225

Ala Arg  Ile Ser Gln Gly Ala  Gly Trp Ser Leu Arg  Glu Thr Ala
    3230                3235                3240

Cys Leu  Gly Lys Ser Tyr Ala  Gln Met Trp Ser Leu  Met Tyr Phe
    3245                3250                3255

His Arg  Arg Asp Leu Arg Leu  Ala Ala Asn Ala Ile  Cys Ser Ala
    3260                3265                3270
```

```
Val Pro  Ser His Trp Val Pro  Thr Ser Arg Thr Thr  Trp Ser Ile
    3275                 3280                 3285

His Ala  Lys His Glu Trp Met  Thr Thr Glu Asp Met  Leu Thr Val
    3290                 3295                 3300

Trp Asn  Arg Val Trp Ile Gln  Glu Asn Pro Trp Met  Glu Asp Lys
    3305                 3310                 3315

Thr Pro  Val Glu Ser Trp Glu  Glu Ile Pro Tyr Leu  Gly Lys Arg
    3320                 3325                 3330

Glu Asp  Gln Trp Cys Gly Ser  Leu Ile Gly Leu Thr  Ser Arg Ala
    3335                 3340                 3345

Thr Trp  Ala Lys Asn Ile Gln  Ala Ala Ile Asn Gln  Val Arg Ser
    3350                 3355                 3360

Leu Ile  Gly Asn Glu Glu Tyr  Thr Asp Tyr Met Pro  Ser Met Lys
    3365                 3370                 3375

Arg Phe  Arg Arg Glu Glu Glu  Glu Ala Gly Val Leu  Trp
    3380                 3385                 3390
```

<210> 33
<211> 3464
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, WVS

<220>
<221> dấu hiệu hỗn tạp
<222> (763)..(763)
<223> Xaa là alanin hoặc glyxin

<400> 33

```
Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1               5                  10                  15
```

```
Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
            20                  25                  30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
        35                  40                  45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
    50                  55                  60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65                  70                  75                  80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
                85                  90                  95

Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
            100                 105                 110

Met Ala Phe His Leu Thr Thr Arg Gly Gly Glu Pro His Met Ile Val
        115                 120                 125

Ser Lys Gln Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly
    130                 135                 140

Val Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Leu Cys Glu
145                 150                 155                 160

Asp Thr Met Thr Tyr Lys Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp
                165                 170                 175

Asp Val Asp Cys Trp Cys Asn Ala Thr Asp Thr Trp Val Thr Tyr Gly
            180                 185                 190

Thr Cys Ser Gln Thr Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
        195                 200                 205

Leu Ala Pro His Val Gly Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp
    210                 215                 220
```

```
Met Ser Ser Glu Gly Ala Trp Lys Gln Ile Gln Lys Val Glu Thr Trp
225                 230                 235                 240

Ala Leu Arg His Pro Gly Phe Thr Val Ile Ala Leu Phe Leu Ala His
                245                 250                 255

Ala Ile Gly Thr Ser Ile Thr Gln Lys Gly Ile Ile Phe Ile Leu Leu
            260                 265                 270

Met Leu Val Thr Pro Ser Met Ala Met Arg Cys Val Gly Ile Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
    290                 295                 300

Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Ile Glu Leu Leu Lys Thr Glu Val Thr Asn Pro Ala Val Leu
                325                 330                 335

Arg Lys Leu Cys Ile Glu Ala Lys Ile Ser Asn Thr Thr Thr Asp Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Ala Thr Leu Val Glu Glu Gln Asp Ala
        355                 360                 365

Asn Phe Val Cys Arg Arg Thr Phe Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Ile Thr Cys Ala Lys Phe Lys
385                 390                 395                 400

Cys Val Thr Lys Leu Glu Gly Lys Ile Val Gln Tyr Glu Asn Leu Lys
                405                 410                 415

Tyr Ser Val Ile Val Thr Val His Thr Gly Asp Gln His Gln Val Gly
            420                 425                 430
```

```
Asn Glu Thr Thr Glu His Gly Thr Thr Ala Thr Ile Thr Pro Gln Ala
        435                 440                 445

Pro Thr Ser Glu Ile Gln Leu Thr Asp Tyr Gly Thr Leu Thr Leu Asp
    450                 455                 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Thr
465                 470                 475                 480

Met Lys Glu Arg Ser Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
                485                 490                 495

Pro Leu Pro Trp Thr Ser Gly Ala Ser Thr Ser Gln Glu Thr Trp Asn
            500                 505                 510

Arg Gln Asp Leu Leu Val Thr Phe Lys Thr Ala His Ala Lys Lys Gln
        515                 520                 525

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

Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Thr Thr Thr Ile Phe Ala
545                 550                 555                 560

Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr Leu Lys Gly
                565                 570                 575

Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys Glu Val
            580                 585                 590

Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr Glu Gly
        595                 600                 605

Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu Lys Gly
    610                 615                 620

Ala Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val Thr Asp
625                 630                 635                 640
```

```
Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
                645                 650                 655

Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser Trp Phe
            660                 665                 670

Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly
        675                 680                 685

Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
    690                 695                 700

Ile Gly Gly Val Phe Thr Ser Met Gly Lys Leu Val His Gln Val Phe
705                 710                 715                 720

Gly Thr Ala Tyr Gly Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
                725                 730                 735

Ile Gly Ile Gly Ile Leu Leu Thr Trp Leu Gly Leu Asn Ser Arg Asn
            740                 745                 750

Thr Ser Leu Ser Met Met Cys Ile Ala Ala Xaa Ile Val Thr Leu Tyr
        755                 760                 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
    770                 775                 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785                 790                 795                 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
                805                 810                 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
            820                 825                 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
        835                 840                 845
```

```
Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
    850                 855                 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865                 870                 875                 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
                885                 890                 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
            900                 905                 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
        915                 920                 925

Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
            980                 985                 990

Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035

Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050
```

```
Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065

Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                1075                1080

Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                1090                1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                1105                1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                1120                1125

His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                1135                1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                1150                1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                1165                1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                1180                1185

Ala Thr  Met Thr Asp Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                1195                1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                1210                1215

Leu Leu  Arg Lys Leu Thr Ser  Lys Glu Leu Met Met  Thr Thr Ile
    1220                1225                1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Leu Pro Glu  Thr Ile Leu
    1235                1240                1245
```

```
Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                 1255                 1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                 1270                 1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                 1285                 1290

Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                 1300                 1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                 1315                 1320

Thr Ile  Lys Gly Leu Asn Pro  Thr Ala Ile Phe Leu  Thr Thr Leu
    1325                 1330                 1335

Ser Arg  Thr Ser Lys Lys Arg  Ser Trp Pro Leu Asn  Glu Ala Ile
    1340                 1345                 1350

Met Ala  Val Gly Met Val Ser  Ile Leu Ala Ser Ser  Leu Leu Lys
    1355                 1360                 1365

Asn Asp  Ile Pro Met Thr Gly  Pro Leu Val Ala Gly  Gly Leu Leu
    1370                 1375                 1380

Thr Val  Cys Tyr Val Leu Thr  Gly Arg Ser Ala Asp  Leu Glu Leu
    1385                 1390                 1395

Glu Arg  Ala Ala Asp Val Lys  Trp Glu Asp Gln Ala  Glu Ile Ser
    1400                 1405                 1410

Gly Ser  Ser Pro Ile Leu Ser  Ile Thr Ile Ser Glu  Asp Gly Ser
    1415                 1420                 1425

Met Ser  Ile Lys Asn Glu Glu  Glu Asp Gln Thr Leu  Thr Ile Leu
    1430                 1435                 1440
```

```
Ile Arg  Thr Gly Leu Leu Val  Ile Ser Gly Leu Phe  Pro Val Ser
    1445                1450                1455

Ile Pro  Ile Thr Ala Ala Ala  Trp Tyr Leu Trp Glu  Val Lys Lys
    1460                1465                1470

Gln Arg  Ala Gly Val Leu Trp  Asp Val Pro Ser Pro  Pro Pro Met
    1475                1480                1485

Gly Lys  Ala Glu Leu Glu Asp  Gly Ala Tyr Arg Ile  Lys Gln Lys
    1490                1495                1500

Gly Ile  Leu Gly Tyr Ser Gln  Ile Gly Ala Gly Val  Tyr Lys Glu
    1505                1510                1515

Gly Thr  Phe His Thr Met Trp  His Val Thr Arg Gly  Ala Val Leu
    1520                1525                1530

Met His  Lys Gly Lys Arg Ile  Glu Pro Ser Trp Ala  Asp Val Lys
    1535                1540                1545

Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                1555                1560

Trp Lys  Glu Gly Glu Glu Val  Gln Val Leu Ala Leu  Glu Pro Gly
    1565                1570                1575

Lys Asn  Pro Arg Ala Val Gln  Thr Lys Pro Gly Leu  Phe Lys Thr
    1580                1585                1590

Asn Ala  Gly Thr Ile Gly Ala  Val Ser Leu Asp Phe  Ser Pro Gly
    1595                1600                1605

Thr Ser  Gly Ser Pro Ile Ile  Asp Lys Lys Gly Lys  Val Val Gly
    1610                1615                1620

Leu Tyr  Gly Asn Gly Val Val  Thr Arg Ser Gly Ala  Tyr Val Ser
    1625                1630                1635
```

```
Ala Ile  Ala Gln Thr Glu Lys  Ser Ile Glu Asp Asn  Pro Glu Ile
    1640                1645                1650

Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                1660                1665

His Pro  Gly Ala Gly Lys Thr  Lys Arg Tyr Leu Pro  Ala Ile Val
    1670                1675                1680

Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                1690                1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                1705                1710

Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                1720                1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                1735                1740

Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                1750                1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                1765                1770

Tyr Ile  Ser Thr Arg Val Glu  Met Gly Glu Ala Ala  Gly Ile Phe
    1775                1780                1785

Met Thr  Ala Thr Pro Pro Gly  Ser Arg Asp Pro Phe  Pro Gln Ser
    1790                1795                1800

Asn Ala  Pro Ile Ile Asp Glu  Glu Arg Glu Ile Pro  Glu Arg Ser
    1805                1810                1815

Trp Asn  Ser Gly His Glu Trp  Val Thr Asp Phe Lys  Gly Lys Thr
    1820                1825                1830
```

```
Val Trp  Phe Val Pro Ser Ile  Lys Ala Gly Asn Asp  Ile Ala Ala
    1835                1840                1845

Cys Leu  Arg Lys Asn Gly Lys  Lys Val Ile Gln Leu  Ser Arg Lys
    1850                1855                1860

Thr Phe  Asp Ser Glu Tyr Val  Lys Thr Arg Thr Asn  Asp Trp Asp
    1865                1870                1875

Phe Val  Val Thr Thr Asp Ile  Ser Glu Met Gly Ala  Asn Phe Lys
    1880                1885                1890

Ala Glu  Arg Val Ile Asp Pro  Arg Arg Cys Met Lys  Pro Val Ile
    1895                1900                1905

Leu Thr  Asp Gly Glu Glu Arg  Val Ile Leu Ala Gly  Pro Met Pro
    1910                1915                1920

Val Thr  His Ser Ser Ala Ala  Gln Arg Arg Gly Arg  Ile Gly Arg
    1925                1930                1935

Asn Pro  Lys Asn Glu Asn Asp  Gln Tyr Ile Tyr Met  Gly Glu Pro
    1940                1945                1950

Leu Glu  Asn Asp Glu Asp Cys  Ala His Trp Lys Glu  Ala Lys Met
    1955                1960                1965

Leu Leu  Asp Asn Ile Asn Thr  Pro Glu Gly Ile Ile  Pro Ser Met
    1970                1975                1980

Phe Glu  Pro Glu Arg Glu Lys  Val Asp Ala Ile Asp  Gly Glu Tyr
    1985                1990                1995

Arg Leu  Arg Gly Glu Ala Arg  Lys Thr Phe Val Asp  Leu Met Arg
    2000                2005                2010

Arg Gly  Asp Leu Pro Val Trp  Leu Ala Tyr Arg Val  Ala Ala Glu
    2015                2020                2025
```

```
Gly Ile  Asn Tyr Ala Asp Arg  Arg Trp Cys Phe Asp  Gly Val Lys
    2030                 2035                 2040

Asn Asn  Gln Ile Leu Glu Glu  Asn Val Glu Val Glu  Ile Trp Thr
    2045                 2050                 2055

Lys Glu  Gly Glu Arg Lys Lys  Leu Lys Pro Arg Trp  Leu Asp Ala
    2060                 2065                 2070

Arg Ile  Tyr Ser Asp Pro Leu  Ala Leu Lys Glu Phe  Lys Glu Phe
    2075                 2080                 2085

Ala Ala  Gly Arg Lys Ser Leu  Thr Leu Asn Leu Ile  Thr Glu Met
    2090                 2095                 2100

Gly Arg  Leu Pro Thr Phe Met  Thr Gln Lys Ala Arg  Asp Ala Leu
    2105                 2110                 2115

Asp Asn  Leu Ala Val Leu His  Thr Ala Glu Ala Gly  Gly Arg Ala
    2120                 2125                 2130

Tyr Asn  His Ala Leu Ser Glu  Leu Pro Glu Thr Leu  Glu Thr Leu
    2135                 2140                 2145

Leu Leu  Leu Thr Leu Leu Ala  Thr Val Thr Gly Gly  Ile Phe Leu
    2150                 2155                 2160

Phe Leu  Met Ser Ala Arg Gly  Ile Gly Lys Met Thr  Leu Gly Met
    2165                 2170                 2175

Cys Cys  Ile Ile Thr Ala Ser  Ile Leu Leu Trp Tyr  Ala Gln Ile
    2180                 2185                 2190

Gln Pro  His Trp Ile Ala Ala  Ser Ile Ile Leu Glu  Phe Phe Leu
    2195                 2200                 2205

Ile Val  Leu Leu Ile Pro Glu  Pro Glu Lys Gln Arg  Thr Pro Gln
    2210                 2215                 2220
```

```
Asp Asn  Gln Leu Thr Tyr Val  Val Ile Ala Ile Leu  Thr Val Val
    2225                 2230                 2235

Ala Ala  Thr Met Ala Asn Glu  Met Gly Phe Leu Glu  Lys Thr Lys
    2240                 2245                 2250

Lys Asp  Leu Gly Leu Gly Ser  Ile Ala Thr Gln Gln  Pro Glu Ser
    2255                 2260                 2265

Asn Ile  Leu Asp Ile Asp Leu  Arg Pro Ala Ser Ala  Trp Thr Leu
    2270                 2275                 2280

Tyr Ala  Val Ala Thr Thr Phe  Val Thr Pro Met Leu  Arg His Ser
    2285                 2290                 2295

Ile Glu  Asn Ser Ser Val Asn  Val Ser Leu Thr Ala  Ile Ala Asn
    2300                 2305                 2310

Gln Ala  Thr Val Leu Met Gly  Leu Gly Lys Gly Trp  Pro Leu Ser
    2315                 2320                 2325

Lys Met  Asp Ile Gly Val Pro  Leu Leu Ala Ile Gly  Cys Tyr Ser
    2330                 2335                 2340

Gln Val  Asn Pro Ile Thr Leu  Ile Ala Ala Leu Phe  Leu Leu Val
    2345                 2350                 2355

Ala His  Tyr Ala Ile Ile Gly  Pro Gly Leu Gln Ala  Lys Ala Thr
    2360                 2365                 2370

Arg Glu  Ala Gln Lys Arg Ala  Ala Ala Gly Ile Met  Lys Asn Pro
    2375                 2380                 2385

Thr Val  Asp Gly Ile Thr Val  Ile Asp Leu Asp Pro  Ile Pro Tyr
    2390                 2395                 2400

Asp Pro  Lys Phe Glu Lys Gln  Leu Gly Gln Val Met  Leu Leu Val
    2405                 2410                 2415
```

```
Leu Cys  Val Thr Gln Val Leu  Met Met Arg Thr Thr  Trp Ala Leu
    2420                 2425                 2430

Cys Glu  Ala Leu Thr Leu Ala  Thr Gly Pro Ile Ser  Thr Leu Trp
    2435                 2440                 2445

Glu Gly  Asn Pro Gly Arg Phe  Trp Asn Thr Thr Ile  Ala Val Ser
    2450                 2455                 2460

Met Ala  Asn Ile Phe Arg Gly  Ser Tyr Leu Ala Gly  Ala Gly Leu
    2465                 2470                 2475

Leu Phe  Ser Ile Met Lys Asn  Thr Thr Asn Thr Arg  Arg Gly Thr
    2480                 2485                 2490

Gly Asn  Ile Gly Glu Thr Leu  Gly Glu Lys Trp Lys  Ser Arg Leu
    2495                 2500                 2505

Asn Ala  Leu Gly Lys Ser Glu  Phe Gln Ile Tyr Lys  Lys Ser Gly
    2510                 2515                 2520

Ile Gln  Glu Val Asp Arg Thr  Leu Ala Lys Glu Gly  Ile Lys Arg
    2525                 2530                 2535

Gly Glu  Thr Asp His His Ala  Val Ser Arg Gly Ser  Ala Lys Leu
    2540                 2545                 2550

Arg Trp  Phe Val Glu Arg Asn  Met Val Thr Pro Glu  Gly Lys Val
    2555                 2560                 2565

Val Asp  Leu Gly Cys Gly Arg  Gly Gly Trp Ser Tyr  Tyr Cys Gly
    2570                 2575                 2580

Gly Leu  Lys Asn Val Arg Glu  Val Lys Gly Leu Thr  Lys Gly Gly
    2585                 2590                 2595

Pro Gly  His Glu Glu Pro Ile  Pro Met Ser Thr Tyr  Gly Trp Asn
    2600                 2605                 2610
```

```
Leu Val  Arg Leu Gln Ser Gly  Val Asp Val Phe Phe  Ile Pro Pro
    2615                 2620                 2625

Glu Lys  Cys Asp Thr Leu Leu  Cys Asp Ile Gly Glu  Ser Ser Pro
    2630                 2635                 2640

Asn Pro  Thr Val Glu Ala Gly  Arg Thr Leu Arg Val  Leu Asn Leu
    2645                 2650                 2655

Val Glu  Asn Trp Leu Asn Asn  Asn Thr Gln Phe Cys  Ile Lys Val
    2660                 2665                 2670

Leu Asn  Pro Tyr Met Pro Ser  Val Ile Glu Lys Met  Glu Ala Leu
    2675                 2680                 2685

Gln Arg  Lys Tyr Gly Gly Ala  Leu Val Arg Asn Pro  Leu Ser Arg
    2690                 2695                 2700

Asn Ser  Thr His Glu Met Tyr  Trp Val Ser Asn Ala  Ser Gly Asn
    2705                 2710                 2715

Ile Val  Ser Ser Val Asn Met  Ile Ser Arg Met Leu  Ile Asn Arg
    2720                 2725                 2730

Phe Thr  Met Arg Tyr Lys Lys  Ala Thr Tyr Glu Pro  Asp Val Asp
    2735                 2740                 2745

Leu Gly  Ser Gly Thr Arg Asn  Ile Gly Ile Glu Ser  Glu Ile Pro
    2750                 2755                 2760

Asn Leu  Asp Ile Ile Gly Lys  Arg Ile Glu Lys Ile  Lys Gln Glu
    2765                 2770                 2775

His Glu  Thr Ser Trp His Tyr  Asp Gln Asp His Pro  Tyr Lys Thr
    2780                 2785                 2790

Trp Ala  Tyr His Gly Ser Tyr  Glu Thr Lys Gln Thr  Gly Ser Ala
    2795                 2800                 2805
```

```
Ser Ser  Met Val Asn Gly Val  Val Arg Leu Leu Thr  Lys Pro Trp
    2810                 2815                 2820

Asp Val  Val Pro Met Val Thr  Gln Met Ala Met Thr  Asp Thr Thr
    2825                 2830                 2835

Pro Phe  Gly Gln Gln Arg Val  Phe Lys Glu Lys Val  Asp Thr Arg
    2840                 2845                 2850

Thr Gln  Glu Pro Lys Glu Gly  Thr Lys Lys Leu Met  Lys Ile Thr
    2855                 2860                 2865

Ala Glu  Trp Leu Trp Lys Glu  Leu Gly Lys Lys Lys  Thr Pro Arg
    2870                 2875                 2880

Met Cys  Thr Arg Glu Glu Phe  Thr Arg Lys Val Arg  Ser Asn Ala
    2885                 2890                 2895

Ala Leu  Gly Ala Ile Phe Thr  Asp Glu Asn Lys Trp  Lys Ser Ala
    2900                 2905                 2910

Arg Glu  Ala Val Glu Asp Ser  Arg Phe Trp Glu Leu  Val Asp Lys
    2915                 2920                 2925

Glu Arg  Asn Leu His Leu Glu  Gly Lys Cys Glu Thr  Cys Val Tyr
    2930                 2935                 2940

Asn Met  Met Gly Lys Arg Glu  Lys Lys Leu Gly Glu  Phe Gly Lys
    2945                 2950                 2955

Ala Lys  Gly Ser Arg Ala Ile  Trp Tyr Met Trp Leu  Gly Ala Arg
    2960                 2965                 2970

Phe Leu  Glu Phe Glu Ala Leu  Gly Phe Leu Asn Glu  Asp His Trp
    2975                 2980                 2985

Phe Ser  Arg Glu Asn Ser Leu  Ser Gly Val Glu Gly  Glu Gly Leu
    2990                 2995                 3000
```

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
3080 3085 3090

Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala
3095 3100 3105

Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile
3110 3115 3120

Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
3125 3130 3135

Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly
3140 3145 3150

Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
3155 3160 3165

Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln
3170 3175 3180

Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro
3185 3190 3195

```
Phe Cys  Ser His His Phe His  Glu Leu Ile Met Lys  Asp Gly Arg
    3200                3205                3210

Val Leu  Val Val Pro Cys Arg  Asn Gln Asp Glu Leu  Ile Gly Arg
    3215                3220                3225

Ala Arg  Ile Ser Gln Gly Ala  Gly Trp Ser Leu Arg  Glu Thr Ala
    3230                3235                3240

Cys Leu  Gly Lys Ser Tyr Ala  Gln Met Trp Ser Leu  Met Tyr Phe
    3245                3250                3255

His Arg  Arg Asp Leu Arg Leu  Ala Ala Asn Ala Ile  Cys Ser Ala
    3260                3265                3270

Val Pro  Ser His Trp Val Pro  Thr Ser Arg Thr Thr  Trp Ser Ile
    3275                3280                3285

His Ala  Lys His Glu Trp Met  Thr Thr Glu Asp Met  Leu Thr Val
    3290                3295                3300

Trp Asn  Arg Val Trp Ile Gln  Glu Asn Pro Trp Met  Glu Asp Lys
    3305                3310                3315

Thr Pro  Val Glu Ser Trp Glu  Glu Ile Pro Tyr Leu  Gly Lys Arg
    3320                3325                3330

Glu Asp  Gln Trp Cys Gly Ser  Leu Ile Gly Leu Thr  Ser Arg Ala
    3335                3340                3345

Thr Trp  Ala Lys Asn Ile Gln  Ala Ala Ile Asn Gln  Val Arg Ser
    3350                3355                3360

Leu Ile  Gly Asn Glu Glu Tyr  Thr Asp Tyr Met Pro  Ser Met Lys
    3365                3370                3375

Arg Phe  Arg Arg Glu Glu Glu  Glu Ala Gly Val Leu  Trp Lys Ala
    3380                3385                3390
```

```
Lys Leu  Thr Asn Lys Ala Arg  Ser Gln Val Gly Leu  Ser His Ser
    3395                 3400                 3405

Thr Glu  Lys Thr Met Leu Pro  Val Ser Pro Val Gln  Gly Arg Lys
    3410                 3415                 3420

Lys Ser  Gly His His Lys Cys  His Ser Leu Ser Lys  Leu Cys Ser
    3425                 3430                 3435

Leu Leu  His Leu Arg Arg Cys  Lys Lys Ser Gly Arg  Pro Gln Thr
    3440                 3445                 3450

Met Glu  Ala Val Arg Met Ala  Trp Thr Ser Gly
    3455                 3460
```

<210> 34
<211> 3464
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Typ huyết thanh virut sốt xuất huyết 1, BVS

<400> 34

```
Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu
1               5                  10                  15

Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg
            20                  25                  30

Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met
        35                  40                  45

Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly
    50                  55                  60

Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val
65                  70                  75                  80

Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn
                85                  90                  95
```

```
Arg Arg Arg Arg Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val
            100                 105                 110

Met Ala Phe His Leu Thr Thr Arg Gly Gly Glu Pro His Met Ile Val
        115                 120                 125

Ser Lys Gln Glu Arg Gly Lys Ser Leu Leu Phe Lys Thr Ser Ala Gly
    130                 135                 140

Val Asn Met Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Leu Cys Glu
145                 150                 155                 160

Asp Thr Met Thr Tyr Lys Cys Pro Arg Ile Thr Glu Ala Glu Pro Asp
                165                 170                 175

Asp Val Asp Cys Trp Cys Asn Ala Thr Asp Thr Trp Val Thr Tyr Gly
            180                 185                 190

Thr Cys Ser Gln Thr Gly Glu His Arg Arg Asp Lys Arg Ser Val Ala
        195                 200                 205

Leu Ala Pro His Val Gly Leu Gly Leu Glu Thr Arg Ala Glu Thr Trp
    210                 215                 220

Met Ser Ser Glu Gly Ala Trp Lys Gln Ile Gln Lys Val Glu Thr Trp
225                 230                 235                 240

Ala Leu Arg His Pro Gly Phe Thr Val Ile Ala Leu Phe Leu Ala His
                245                 250                 255

Ala Ile Gly Thr Ser Ile Thr Gln Lys Gly Ile Ile Phe Ile Leu Leu
            260                 265                 270

Met Leu Val Thr Pro Ser Met Ala Met Arg Cys Val Gly Ile Gly Asn
        275                 280                 285

Arg Asp Phe Val Glu Gly Leu Ser Gly Ala Thr Trp Val Asp Val Val
    290                 295                 300
```

```
Leu Glu His Gly Ser Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr
305                 310                 315                 320

Leu Asp Ile Glu Leu Leu Lys Thr Glu Val Thr Asn Pro Ala Val Leu
                325                 330                 335

Arg Lys Leu Cys Ile Glu Ala Lys Ile Ser Asn Thr Thr Thr Asp Ser
            340                 345                 350

Arg Cys Pro Thr Gln Gly Glu Ala Thr Leu Val Glu Glu Gln Asp Ala
        355                 360                 365

Asn Phe Val Cys Arg Arg Thr Phe Val Asp Arg Gly Trp Gly Asn Gly
    370                 375                 380

Cys Gly Leu Phe Gly Lys Gly Ser Leu Ile Thr Cys Ala Lys Phe Lys
385                 390                 395                 400

Cys Val Thr Lys Leu Glu Gly Lys Ile Val Gln Tyr Glu Asn Leu Lys
                405                 410                 415

Tyr Ser Val Ile Val Thr Val His Thr Gly Asp Gln His Gln Val Gly
            420                 425                 430

Asn Glu Thr Thr Glu His Gly Thr Thr Ala Thr Ile Thr Pro Gln Ala
        435                 440                 445

Pro Thr Ser Glu Ile Gln Leu Thr Asp Tyr Gly Thr Leu Thr Leu Asp
    450                 455                 460

Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Val Leu Leu Thr
465                 470                 475                 480

Met Lys Glu Arg Ser Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu
                485                 490                 495

Pro Leu Pro Trp Thr Ser Gly Ala Ser Thr Ser Gln Glu Thr Trp Asn
            500                 505                 510
```

```
Arg Gln Asp Leu Leu Val Thr Phe Lys Thr Ala His Ala Lys Lys Gln
        515                 520                 525

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

Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Thr Thr Thr Ile Phe Ala
545                 550                 555                 560

Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Thr Leu Lys Gly
                565                 570                 575

Met Ser Tyr Val Met Cys Thr Gly Ser Phe Lys Leu Glu Lys Glu Val
            580                 585                 590

Ala Glu Thr Gln His Gly Thr Val Leu Val Gln Val Lys Tyr Glu Gly
        595                 600                 605

Thr Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Gln Asp Glu Lys Gly
    610                 615                 620

Ala Thr Gln Asn Gly Arg Leu Ile Thr Ala Asn Pro Ile Val Thr Asp
625                 630                 635                 640

Lys Glu Lys Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser
                645                 650                 655

Tyr Ile Val Val Gly Ala Gly Glu Lys Ala Leu Lys Leu Ser Trp Phe
            660                 665                 670

Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly
        675                 680                 685

Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
    690                 695                 700

Ile Gly Gly Val Phe Thr Ser Met Gly Lys Leu Val His Gln Val Phe
705                 710                 715                 720
```

```
Gly Thr Ala Tyr Gly Val Leu Phe Ser Gly Val Ser Trp Thr Met Lys
                725                 730                 735

Ile Gly Ile Gly Ile Leu Leu Thr Trp Leu Gly Leu Asn Ser Arg Asn
            740                 745                 750

Thr Ser Leu Ser Met Met Cys Ile Ala Ala Ala Ile Val Thr Leu Tyr
        755                 760                 765

Leu Gly Val Met Val Gln Ala Asp Ser Gly Cys Val Val Ser Trp Lys
    770                 775                 780

Asn Lys Glu Leu Lys Cys Gly Ser Gly Ile Phe Ile Thr Asp Asn Val
785                 790                 795                 800

His Thr Trp Thr Glu Gln Tyr Lys Phe Gln Pro Glu Ser Pro Ser Lys
                805                 810                 815

Leu Ala Ser Ala Ile Gln Lys Ala His Glu Glu Asp Ile Cys Gly Ile
            820                 825                 830

Arg Ser Val Thr Arg Leu Glu Asn Leu Met Trp Lys Gln Ile Thr Pro
        835                 840                 845

Glu Leu Asn His Ile Leu Ser Glu Asn Glu Val Lys Leu Thr Ile Met
    850                 855                 860

Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg
865                 870                 875                 880

Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala
                885                 890                 895

Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly
            900                 905                 910

Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu
        915                 920                 925
```

```
Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu
    930                 935                 940

Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser
945                 950                 955                 960

Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp
                965                 970                 975

Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe
            980                 985                 990

Ile Glu Val Lys Asn Cys His Trp  Pro Lys Ser His Thr  Leu Trp Ser
        995                 1000                1005

Asn Gly  Val Leu Glu Ser Glu  Met Ile Ile Pro Lys  Asn Leu Ala
    1010                1015                1020

Gly Pro  Val Ser Gln His Asn  Tyr Arg Pro Gly Tyr  His Thr Gln
    1025                1030                1035

Ile Thr  Gly Pro Trp His Leu  Gly Lys Leu Glu Met  Asp Phe Asp
    1040                1045                1050

Phe Cys  Asp Gly Thr Thr Val  Val Val Thr Glu Asp  Cys Gly Asn
    1055                1060                1065

Arg Gly  Pro Ser Leu Arg Thr  Thr Thr Ala Ser Gly  Lys Leu Ile
    1070                1075                1080

Thr Glu  Trp Cys Cys Arg Ser  Cys Thr Leu Pro Pro  Leu Arg Tyr
    1085                1090                1095

Arg Gly  Glu Asp Gly Cys Trp  Tyr Gly Met Glu Ile  Arg Pro Leu
    1100                1105                1110

Lys Glu  Lys Glu Glu Asn Leu  Val Asn Ser Leu Val  Thr Ala Gly
    1115                1120                1125
```

```
His Gly  Gln Val Asp Asn Phe  Ser Leu Gly Val Leu  Gly Met Ala
    1130                1135                1140

Leu Phe  Leu Glu Glu Met Leu  Arg Thr Arg Val Gly  Thr Lys His
    1145                1150                1155

Ala Ile  Leu Leu Val Ala Val  Ser Phe Val Thr Leu  Ile Thr Gly
    1160                1165                1170

Asn Met  Ser Phe Arg Asp Leu  Gly Arg Val Met Val  Met Val Gly
    1175                1180                1185

Ala Thr  Met Thr Asp Asp Ile  Gly Met Gly Val Thr  Tyr Leu Ala
    1190                1195                1200

Leu Leu  Ala Ala Phe Lys Val  Arg Pro Thr Phe Ala  Ala Gly Leu
    1205                1210                1215

Leu Leu  Arg Lys Leu Thr Ser  Lys Glu Leu Met Met  Thr Thr Ile
    1220                1225                1230

Gly Ile  Val Leu Leu Ser Gln  Ser Thr Leu Pro Glu  Thr Ile Leu
    1235                1240                1245

Glu Leu  Thr Asp Ala Leu Ala  Leu Gly Met Met Val  Leu Lys Met
    1250                1255                1260

Val Arg  Asn Met Glu Lys Tyr  Gln Leu Ala Val Thr  Ile Met Ala
    1265                1270                1275

Ile Leu  Cys Val Pro Asn Ala  Val Ile Leu Gln Asn  Ala Trp Lys
    1280                1285                1290

Val Ser  Cys Thr Ile Leu Ala  Val Val Ser Val Ser  Pro Leu Phe
    1295                1300                1305

Leu Thr  Ser Ser Gln Gln Lys  Thr Asp Trp Ile Pro  Leu Ala Leu
    1310                1315                1320
```

```
Thr Ile  Lys Gly Leu Asn Pro  Thr Ala Ile Phe Leu  Thr Thr Leu
    1325                1330                1335

Ser Arg  Thr Ser Lys Lys Arg  Ser Trp Pro Leu Asn  Glu Ala Ile
    1340                1345                1350

Met Ala  Val Gly Met Val Ser  Ile Leu Ala Ser Ser  Leu Leu Lys
    1355                1360                1365

Asn Asp  Ile Pro Met Thr Gly  Pro Leu Val Ala Gly  Gly Leu Leu
    1370                1375                1380

Thr Val  Cys Tyr Val Leu Thr  Gly Arg Ser Ala Asp  Leu Glu Leu
    1385                1390                1395

Glu Arg  Ala Ala Asp Val Lys  Trp Glu Asp Gln Ala  Glu Ile Ser
    1400                1405                1410

Gly Ser  Ser Pro Ile Leu Ser  Ile Thr Ile Ser Glu  Asp Gly Ser
    1415                1420                1425

Met Ser  Ile Lys Asn Glu Glu  Glu Asp Gln Thr Leu  Thr Ile Leu
    1430                1435                1440

Ile Arg  Thr Gly Leu Leu Val  Ile Ser Gly Leu Phe  Pro Val Ser
    1445                1450                1455

Ile Pro  Ile Thr Ala Ala Ala  Trp Tyr Leu Trp Glu  Val Lys Lys
    1460                1465                1470

Gln Arg  Ala Gly Val Leu Trp  Asp Val Pro Ser Pro  Pro Pro Met
    1475                1480                1485

Gly Lys  Ala Glu Leu Glu Asp  Gly Ala Tyr Arg Ile  Lys Gln Lys
    1490                1495                1500

Gly Ile  Leu Gly Tyr Ser Gln  Ile Gly Ala Gly Val  Tyr Lys Glu
    1505                1510                1515
```

```
Gly Thr  Phe His Thr Met Trp  His Val Thr Arg Gly  Ala Val Leu
    1520                1525                1530

Met His  Lys Gly Lys Arg Ile  Glu Pro Ser Trp Ala  Asp Val Lys
    1535                1540                1545

Lys Asp  Leu Ile Ser Tyr Gly  Gly Gly Trp Lys Leu  Glu Gly Glu
    1550                1555                1560

Trp Lys  Glu Gly Glu Glu Val  Gln Val Leu Ala Leu  Glu Pro Gly
    1565                1570                1575

Lys Asn  Pro Arg Ala Val Gln  Thr Lys Pro Gly Leu  Phe Lys Thr
    1580                1585                1590

Asn Ala  Gly Thr Ile Gly Ala  Val Ser Leu Asp Phe  Ser Pro Gly
    1595                1600                1605

Thr Ser  Gly Ser Pro Ile Ile  Asp Lys Lys Gly Lys  Val Val Gly
    1610                1615                1620

Leu Tyr  Gly Asn Gly Val Val  Thr Arg Ser Gly Ala  Tyr Val Ser
    1625                1630                1635

Ala Ile  Ala Gln Thr Glu Lys  Ser Ile Glu Asp Asn  Pro Glu Ile
    1640                1645                1650

Glu Asp  Asp Ile Phe Arg Lys  Arg Arg Leu Thr Ile  Met Asp Leu
    1655                1660                1665

His Pro  Gly Ala Gly Lys Thr  Lys Arg Tyr Leu Pro  Ala Ile Val
    1670                1675                1680

Arg Glu  Ala Ile Lys Arg Gly  Leu Arg Thr Leu Ile  Leu Ala Pro
    1685                1690                1695

Thr Arg  Val Val Ala Ala Glu  Met Glu Glu Ala Leu  Arg Gly Leu
    1700                1705                1710
```

```
Pro Ile  Arg Tyr Gln Thr Pro  Ala Ile Arg Ala Val  His Thr Gly
    1715                 1720                 1725

Arg Glu  Ile Val Asp Leu Met  Cys His Ala Thr Phe  Thr Met Arg
    1730                 1735                 1740

Leu Leu  Ser Pro Val Arg Val  Pro Asn Tyr Asn Leu  Ile Ile Met
    1745                 1750                 1755

Asp Glu  Ala His Phe Thr Asp  Pro Ala Ser Ile Ala  Ala Arg Gly
    1760                 1765                 1770

Tyr Ile  Ser Thr Arg Val Glu  Met Gly Glu Ala Ala  Gly Ile Phe
    1775                 1780                 1785

Met Thr  Ala Thr Pro Pro Gly  Ser Arg Asp Pro Phe  Pro Gln Ser
    1790                 1795                 1800

Asn Ala  Pro Ile Ile Asp Glu  Glu Arg Glu Ile Pro  Glu Arg Ser
    1805                 1810                 1815

Trp Asn  Ser Gly His Glu Trp  Val Thr Asp Phe Lys  Gly Lys Thr
    1820                 1825                 1830

Val Trp  Phe Val Pro Ser Ile  Lys Ala Gly Asn Asp  Ile Ala Ala
    1835                 1840                 1845

Cys Leu  Arg Lys Asn Gly Lys  Lys Val Ile Gln Leu  Ser Arg Lys
    1850                 1855                 1860

Thr Phe  Asp Ser Glu Tyr Val  Lys Thr Arg Thr Asn  Asp Trp Asp
    1865                 1870                 1875

Phe Val  Val Thr Thr Asp Ile  Ser Glu Met Gly Ala  Asn Phe Lys
    1880                 1885                 1890

Ala Glu  Arg Val Ile Asp Pro  Arg Arg Cys Met Lys  Pro Val Ile
    1895                 1900                 1905
```

Leu Thr Asp Gly Glu Glu Arg Val Ile Leu Ala Gly Pro Met Pro
1910 1915 1920

Val Thr His Ser Ser Ala Ala Gln Arg Arg Gly Arg Ile Gly Arg
1925 1930 1935

Asn Pro Lys Asn Glu Asn Asp Gln Tyr Ile Tyr Met Gly Glu Pro
1940 1945 1950

Leu Glu Asn Asp Glu Asp Cys Ala His Trp Lys Glu Ala Lys Met
1955 1960 1965

Leu Leu Asp Asn Ile Asn Thr Pro Glu Gly Ile Ile Pro Ser Met
1970 1975 1980

Phe Glu Pro Glu Arg Glu Lys Val Asp Ala Ile Asp Gly Glu Tyr
1985 1990 1995

Arg Leu Arg Gly Glu Ala Arg Lys Thr Phe Val Asp Leu Met Arg
2000 2005 2010

Arg Gly Asp Leu Pro Val Trp Leu Ala Tyr Arg Val Ala Ala Glu
2015 2020 2025

Gly Ile Asn Tyr Ala Asp Arg Arg Trp Cys Phe Asp Gly Val Lys
2030 2035 2040

Asn Asn Gln Ile Leu Glu Glu Asn Val Glu Val Glu Ile Trp Thr
2045 2050 2055

Lys Glu Gly Glu Arg Lys Lys Leu Lys Pro Arg Trp Leu Asp Ala
2060 2065 2070

Arg Ile Tyr Ser Asp Pro Leu Ala Leu Lys Glu Phe Lys Glu Phe
2075 2080 2085

Ala Ala Gly Arg Lys Ser Leu Thr Leu Asn Leu Ile Thr Glu Met
2090 2095 2100

Gly Arg Leu Pro Thr Phe Met Thr Gln Lys Ala Arg Asp Ala Leu
2105 2110 2115

Asp Asn Leu Ala Val Leu His Thr Ala Glu Ala Gly Gly Arg Ala
2120 2125 2130

Tyr Asn His Ala Leu Ser Glu Leu Pro Glu Thr Leu Glu Thr Leu
2135 2140 2145

Leu Leu Leu Thr Leu Leu Ala Thr Val Thr Gly Gly Ile Phe Leu
2150 2155 2160

Phe Leu Met Ser Ala Arg Gly Ile Gly Lys Met Thr Leu Gly Met
2165 2170 2175

Cys Cys Ile Ile Thr Ala Ser Ile Leu Leu Trp Tyr Ala Gln Ile
2180 2185 2190

Gln Pro His Trp Ile Ala Ala Ser Ile Ile Leu Glu Phe Phe Leu
2195 2200 2205

Ile Val Leu Leu Ile Pro Glu Pro Glu Lys Gln Arg Thr Pro Gln
2210 2215 2220

Asp Asn Gln Leu Thr Tyr Val Val Ile Ala Ile Leu Thr Val Val
2225 2230 2235

Ala Ala Thr Met Ala Asn Glu Met Gly Phe Leu Glu Lys Thr Lys
2240 2245 2250

Lys Asp Leu Gly Leu Gly Ser Ile Ala Thr Gln Gln Pro Glu Ser
2255 2260 2265

Asn Ile Leu Asp Ile Asp Leu Arg Pro Ala Ser Ala Trp Thr Leu
2270 2275 2280

Tyr Ala Val Ala Thr Thr Phe Val Thr Pro Met Leu Arg His Ser
2285 2290 2295

Ile Glu Asn Ser Ser Val Asn Val Ser Leu Thr Ala Ile Ala Asn
2300 2305 2310

Gln Ala Thr Val Leu Met Gly Leu Gly Lys Gly Trp Pro Leu Ser
2315 2320 2325

Lys Met Asp Ile Gly Val Pro Leu Leu Ala Ile Gly Cys Tyr Ser
2330 2335 2340

Gln Val Asn Pro Ile Thr Leu Ile Ala Ala Leu Phe Leu Leu Val
2345 2350 2355

Ala His Tyr Ala Ile Ile Gly Pro Gly Leu Gln Ala Lys Ala Thr
2360 2365 2370

Arg Glu Ala Gln Lys Arg Ala Ala Ala Gly Ile Met Lys Asn Pro
2375 2380 2385

Thr Val Asp Gly Ile Thr Val Ile Asp Leu Asp Pro Ile Pro Tyr
2390 2395 2400

Asp Pro Lys Phe Glu Lys Gln Leu Gly Gln Val Met Leu Leu Val
2405 2410 2415

Leu Cys Val Thr Gln Val Leu Met Met Arg Thr Thr Trp Ala Leu
2420 2425 2430

Cys Glu Ala Leu Thr Leu Ala Thr Gly Pro Ile Ser Thr Leu Trp
2435 2440 2445

Glu Gly Asn Pro Gly Arg Phe Trp Asn Thr Thr Ile Ala Val Ser
2450 2455 2460

Met Ala Asn Ile Phe Arg Gly Ser Tyr Leu Ala Gly Ala Gly Leu
2465 2470 2475

Leu Phe Ser Ile Met Lys Asn Thr Thr Asn Thr Arg Arg Gly Thr
2480 2485 2490

Gly Asn Ile Gly Glu Thr Leu Gly Glu Lys Trp Lys Ser Arg Leu
2495 2500 2505

Asn Ala Leu Gly Lys Ser Glu Phe Gln Ile Tyr Lys Lys Ser Gly
2510 2515 2520

Ile Gln Glu Val Asp Arg Thr Leu Ala Lys Glu Gly Ile Lys Arg
2525 2530 2535

Gly Glu Thr Asp His His Ala Val Ser Arg Gly Ser Ala Lys Leu
2540 2545 2550

Arg Trp Phe Val Glu Arg Asn Met Val Thr Pro Glu Gly Lys Val
2555 2560 2565

Val Asp Leu Gly Cys Gly Arg Gly Gly Trp Ser Tyr Tyr Cys Gly
2570 2575 2580

Gly Leu Lys Asn Val Arg Glu Val Lys Gly Leu Thr Lys Gly Gly
2585 2590 2595

Pro Gly His Glu Glu Pro Ile Pro Met Ser Thr Tyr Gly Trp Asn
2600 2605 2610

Leu Val Arg Leu Gln Ser Gly Val Asp Val Phe Phe Ile Pro Pro
2615 2620 2625

Glu Lys Cys Asp Thr Leu Leu Cys Asp Ile Gly Glu Ser Ser Pro
2630 2635 2640

Asn Pro Thr Val Glu Ala Gly Arg Thr Leu Arg Val Leu Asn Leu
2645 2650 2655

Val Glu Asn Trp Leu Asn Asn Asn Thr Gln Phe Cys Ile Lys Val
2660 2665 2670

Leu Asn Pro Tyr Met Pro Ser Val Ile Glu Lys Met Glu Ala Leu
2675 2680 2685

Gln Arg Lys Tyr Gly Gly Ala Leu Val Arg Asn Pro Leu Ser Arg
2690 2695 2700

Asn Ser Thr His Glu Met Tyr Trp Val Ser Asn Ala Ser Gly Asn
2705 2710 2715

Ile Val Ser Ser Val Asn Met Ile Ser Arg Met Leu Ile Asn Arg
2720 2725 2730

Phe Thr Met Arg Tyr Lys Lys Ala Thr Tyr Glu Pro Asp Val Asp
2735 2740 2745

Leu Gly Ser Gly Thr Arg Asn Ile Gly Ile Glu Ser Glu Ile Pro
2750 2755 2760

Asn Leu Asp Ile Ile Gly Lys Arg Ile Glu Lys Ile Lys Gln Glu
2765 2770 2775

His Glu Thr Ser Trp His Tyr Asp Gln Asp His Pro Tyr Lys Thr
2780 2785 2790

Trp Ala Tyr His Gly Ser Tyr Glu Thr Lys Gln Thr Gly Ser Ala
2795 2800 2805

Ser Ser Met Val Asn Gly Val Val Arg Leu Leu Thr Lys Pro Trp
2810 2815 2820

Asp Val Val Pro Met Val Thr Gln Met Ala Met Thr Asp Thr Thr
2825 2830 2835

Pro Phe Gly Gln Gln Arg Val Phe Lys Glu Lys Val Asp Thr Arg
2840 2845 2850

Thr Gln Glu Pro Lys Glu Gly Thr Lys Lys Leu Met Lys Ile Thr
2855 2860 2865

Ala Glu Trp Leu Trp Lys Glu Leu Gly Lys Lys Lys Thr Pro Arg
2870 2875 2880

Met Cys Thr Arg Glu Glu Phe Thr Arg Lys Val Arg Ser Asn Ala
2885 2890 2895

Ala Leu Gly Ala Ile Phe Thr Asp Glu Asn Lys Trp Lys Ser Ala
2900 2905 2910

Arg Glu Ala Val Glu Asp Ser Arg Phe Trp Glu Leu Val Asp Lys
2915 2920 2925

Glu Arg Asn Leu His Leu Glu Gly Lys Cys Glu Thr Cys Val Tyr
2930 2935 2940

Asn Met Met Gly Lys Arg Glu Lys Lys Leu Gly Glu Phe Gly Lys
2945 2950 2955

Ala Lys Gly Ser Arg Ala Ile Trp Tyr Met Trp Leu Gly Ala Arg
2960 2965 2970

Phe Leu Glu Phe Glu Ala Leu Gly Phe Leu Asn Glu Asp His Trp
2975 2980 2985

Phe Ser Arg Glu Asn Ser Leu Ser Gly Val Glu Gly Glu Gly Leu
2990 2995 3000

His Lys Leu Gly Tyr Ile Leu Arg Asp Val Ser Lys Lys Glu Gly
3005 3010 3015

Gly Ala Met Tyr Ala Asp Asp Thr Ala Gly Trp Asp Thr Arg Ile
3020 3025 3030

Thr Leu Glu Asp Leu Lys Asn Glu Glu Met Val Thr Asn His Met
3035 3040 3045

Glu Gly Glu His Lys Lys Leu Ala Glu Ala Ile Phe Lys Leu Thr
3050 3055 3060

Tyr Gln Asn Lys Val Val Arg Val Gln Arg Pro Thr Pro Arg Gly
3065 3070 3075

Thr Val Met Asp Ile Ile Ser Arg Arg Asp Gln Arg Gly Ser Gly
3080 3085 3090

Gln Val Gly Thr Tyr Gly Leu Asn Thr Phe Thr Asn Met Glu Ala
3095 3100 3105

Gln Leu Ile Arg Gln Met Glu Gly Glu Gly Val Phe Lys Ser Ile
3110 3115 3120

Gln His Leu Thr Ile Thr Glu Glu Ile Ala Val Gln Asn Trp Leu
3125 3130 3135

Ala Arg Val Gly Arg Glu Arg Leu Ser Arg Met Ala Ile Ser Gly
3140 3145 3150

Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg Phe Ala Ser Ala
3155 3160 3165

Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg Lys Asp Ile Gln
3170 3175 3180

Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp Thr Gln Val Pro
3185 3190 3195

Phe Cys Ser His His Phe His Glu Leu Ile Met Lys Asp Gly Arg
3200 3205 3210

Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu Leu Ile Gly Arg
3215 3220 3225

Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu Arg Glu Thr Ala
3230 3235 3240

Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser Leu Met Tyr Phe
3245 3250 3255

His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala Ile Cys Ser Ala
3260 3265 3270

Val Pro Ser His Trp Val Pro Thr Ser Arg Thr Thr Trp Ser Ile
3275 3280 3285

His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu Thr Val
3290 3295 3300

Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu Asp Lys
3305 3310 3315

Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly Lys Arg
3320 3325 3330

Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser Arg Ala
3335 3340 3345

Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val Arg Ser
3350 3355 3360

Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser Met Lys
3365 3370 3375

Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val Leu Trp Lys Ala
3380 3385 3390

Lys Leu Thr Asn Lys Ala Arg Ser Gln Val Gly Leu Ser His Ser
3395 3400 3405

Thr Glu Lys Thr Met Leu Pro Val Ser Pro Val Gln Gly Arg Lys
3410 3415 3420

Lys Ser Gly His His Lys Cys His Ser Leu Ser Lys Leu Cys Ser
3425 3430 3435

Leu Leu His Leu Arg Arg Cys Lys Lys Ser Gly Arg Pro Gln Thr
3440 3445 3450

Met Glu Ala Val Arg Met Ala Trp Thr Ser Gly
3455 3460