(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) **1-0033423**

(51)[2020.01] **A61K 48/00**; C12N 15/861

(13) **B**

(21) 1-2020-02147
(22) 19/09/2018
(86) PCT/US2018/051812 19/09/2018
(87) WO/2019/060454 28/03/2019
(30) 62/560,901 20/09/2017 US
(45) 26/09/2022 414
(43) 25/08/2020 389ASC
(73) 4D MOLECULAR THERAPEUTICS INC. (US)
5980 Horton Street, Suite 460, Emeryville, CA 94608, U.S.A.
(72) KIRN, David, H. (US); KOTTERMAN, Melissa (US); SCHAFFER, David (US).
(74) Văn phòng Luật sư Ân Nam (ANNAM IP & LAW)

(54) PROTEIN CAPSIT VIRUT ADENO KẾT HỢP BIẾN THỂ VÀ VIRION VIRUT ADENO KẾT HỢP TÁI TỔ HỢP LÂY NHIỄM

(57) Sáng chế đề cập đến protein capsit virut adeno kết hợp (AAV) biến thể có một hoặc nhiều biến đổi trong trình tự axit amin so với protein capsit AAV mẹ, mà khi có mặt trong virion AAV, khiến cho khả năng lây nhiễm vào trong một hoặc nhiều tế bào cơ tăng lên so với khả năng lây nhiễm vào tế bào cơ bởi virion chứa protein capsit AAV mẹ không được biến đổi. Cũng được đề cập đến trong bản mô tả này là virion AAV tái tổ hợp lây nhiễm và dược phẩm chứa virion rAAV như được mô tả trong bản mô tả này.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập chung đến hạt virut của virut adeno kết hợp (AAV) bao gồm các protein capsit biến thể và đề cập đến việc tạo ra các capsit biến thể này bằng cách sử dụng kỹ thuật tiến hóa trực hệ.

Tình trạng kỹ thuật của sáng chế

Cơ liên quan đến hàng loạt các rối loạn di truyền nghiêm trọng. Cơ là mô đích trong liệu pháp gen dùng cho nhiều bệnh loạn dưỡng cơ và cũng có thể được khai thác như một nhà máy sinh học để sản xuất các yếu tố bài tiết để điều trị bệnh toàn thân. Vận chuyển các gen trị liệu cho mô cơ ở người được coi là nhu cầu cấp thiết khẩn cấp nhất trong việc điều trị các bệnh liên quan đến cơ.

Một cách tiếp cận để thực hiện việc vận chuyển gen trực tiếp đến cơ là trị liệu qua trung gian virut adeno kết hợp (AAV) trên cơ sở gen, trong đó virut adeno kết hợp tái tổ hợp (rAAV) được sử dụng để vận chuyển gen đến một hoặc nhiều tế bào cơ, ví dụ để thay thế gen bị thiếu, để sửa chữa gen khiếm khuyết trội, hoặc để tạo thành khuôn dùng cho trị liệu protein liên tục. Mặc dù liệu pháp gen lâm sàng trên cơ sở AAV ngày càng thành công, nhưng nó vẫn còn nhiều thiết sót liên quan đến đặc tính của vectơ virut, bao gồm, ví dụ, tính hướng đích đến các tế bào cơ mong muốn với hiệu suất cao. Theo đó, trong lĩnh vực kỹ thuật, cần có biến thể AAV mới có khả năng tải nạp vượt trội sẽ gicho việc vận chuyển trên cơ sở gen đến tế bào cơ hiệu quả hơn để điều trị bệnh. Trong lĩnh vực kỹ thuật, cần có các biến thể AAV như vậy, thể hiện đặc tính tải nạp cơ được tăng cường – trong một số trường hợp bao quát, trong các trường hợp khác thích hợp với các loại tế bào cơ nhất định - so với AAV kiểu dại và các biến thể AAV đã biết trong lĩnh vực kỹ thuật.

AAV có trong tự nhiên là virut ADN sợi đơn có chứa ba khung đọc mở, *rep*, *cap,* và *aap*. Gen thứ nhất, *rep*, mã hóa cho bốn protein cần thiết cho sự sao chép của gen (Rep78, Rep68, Rep52, và Rep40), gen thứ hai, *cap*, biểu hiện ba protein cấu trúc (VP1-3) cái mà tập hợp lại để tạo thành capsit virut, và gen thứ ba biểu hiện protein kích hoạt lắp ráp (assembly activating protein - AAP) cần thiết cho quá trình lắp ráp

capsit, chẳng hạn như virut adeno hoặc virut herpes, để kích hoạt việc sao chép. Khi vắng mặt virut trợ giúp, AAV thiết lập trạng thái ngầm trong đó hệ gen của nó được giữ ở thể bổ sung hoặc được tích hợp vào trong nhiễm sắc thể chủ trong locut AAVS1.

Các kỹ thuật tiến hóa trực hệ *in vitro* và *in vivo* có thể được sử dụng để chọn lọc các biến thể AAV mà mang lại sự cải tiến so với các vectơ vận chuyển gen trên cơ sở AAV hiện có. Các kỹ thuật tiến hóa trực hệ này đã được biết đến trong lĩnh vực kỹ thuật và được mô tả, ví dụ, trong công bố đơn PCT WO 2014/194132 và Kotterman & Schaffer (Nature Review Genetics, AOP, được đăng trực tuyến vào ngày 20/5/2014; doi: 10.1038/nrg3742), cả hai tài liệu này đều được kết hợp vào đây bằng cách tham chiếu đến toàn bộ. Tiến hóa trực hệ là phương pháp thao tác di truyền trên capsit cạnh tranh với tiến hóa tự nhiên qua các vòng đa dạng hóa di truyền lặp lại và các quy trình chọn lọc, nhờ đó có thể tích lũy được các đột biến có lợi cải thiện được dần dần chức năng của sinh học phân tử như virion trên cơ sở AAV. Theo cách tiếp cận này, các gen *cap* AAV kiểu dại được đa dạng hóa để tạo ra các thư viện gen lớn được đóng gói để tạo ra thư viện của hạt virut, và áp lực chọn lọc được đặt để phân lập các biến thể độc nhất với kiểu hình vượt trội có thể vượt qua rào cản vận chuyển gen.

Các biến thể AAV đã được bộc lộ trong, ví dụ, bằng sáng chế mỹ số 9,193,956; 9;186;419; 8,632,764; 8,663,624; 8,927,514; 8,628,966; 8,263,396; 8,734,809; 8,889,641; 8,632,764; 8,691,948; 8,299,295; 8,802,440; 8,445,267; 8,906,307; 8,574,583; 8,067,015; 7,588,772; 7,867,484; 8,163,543; 8,283,151; 8,999,678; 7,892,809; 7,906,111; 7,259,151; 7,629,322; 7,220,577; 8,802,080; 7,198,951; 8,318,480; 8,962,332; 7,790,449; 7,282,199; 8,906,675; 8,524,446; 7,712,893; 6,491,907; 8,637,255; 7,186,522; 7,105,345; 6,759,237; 6,984,517; 6,962,815; 7,749,492; 7,259,151; và 6,156,303; Công bố đơn sáng chế Mỹ số 2013/0295614; 2015/0065562; 2014/0364338; 2013/0323226; 2014/0359799; 2013/0059732; 2014/0037585; 2014/0056854; 2013/0296409; 2014/0335054 2013/0195801; 2012/0070899; 2011/0275529; 2011/0171262; 2009/0215879; 2010/0297177; 2010/0203083; 2009/0317417; 2009/0202490; 2012/0220492; 2006/0292117; và 2004/0002159; Công bố châu Âu số 2692731 A1; 2383346 B1; 2359865 B1; 2359866 B1; 2359867 B1; và 2357010 B1; 1791858 B1; 1668143 B1; 1660678 B1; 1664314 B1; 1496944 B1; 1456383 B1; 2341068 B1; 2338900 B1; 1456419 B1; 1310571 B1;

1456383 B1; 1633772 B1; và 1135468 B1; và Công bố quốc tế số WO 2014/124282; WO 2013/170078; WO 2014/160092; WO 2014/103957; WO 2014/052789; WO 2013/174760; WO 2013/123503; WO 2011/038187; và WO 2008/124015; WO 2003/054197; tuy nhiên, không tài liệu nào trong số các tài liệu này mô tả phương án và/hoặc dấu hiệu và/hoặc thành phần của cấu trúc vật chất của biến thể AAV được bộc lộ trong bản mô tả này.

Tất cả các tài liệu và tài liệu tham khảo được đề cập trong bản mô tả này và trong các tài liệu có bằng sáng chế tham khảo đều được đưa vào bản mô tả này bằng cách viện dẫn.

Bản chất kỹ thuật của sáng chế

Được đề xuất trong bản mô tả này là protein capsit của biến thể virut adeno kết hợp (AAV) có một hoặc nhiều biến đổi ở trình tự axit amin so với protein capsit của AAV mẹ, cái mà, khi có mặt trong virion AAV, khiến cho khả năng lây nhiễm vào một hoặc nhiều loại tế bào cơ tăng lên so với khả năng lây nhiễm vào tế bào cơ bởi virion AAV chứa protein capsit AAV mẹ không được biến đổi. Cũng được đề xuất là virion AAV tái tổ hợp và dược phẩm của nó bao gồm protein capsit của AAV biến thể như được mô tả trong bản mô tả này, phương pháp tạo ra protein capsit của rAAV biến thể và virion, và phương pháp sử dụng protein capsit rAAV và hạt virut này trong nghiên cứu và thực hành lâm sàng, ví dụ trong vận chuyển trình tự axit nucleic đến một hoặc nhiều tế bào cơ để điều trị bệnh và rối loạn.

Theo một vài khía cạnh của sáng chế, sáng chế đề xuất protein capsit của virut adeno kết hợp (AAV) biến thể, các protein capsit AAV biến thể này có một hoặc nhiều biến đổi trong trình tự axit amin so với capsit AAV mẹ, cái mà, khi có mặt trong virion AAV, làm cho khả năng lây nhiễm vào một hoặc nhiều tế loại tế bào cơ (ví dụ, tế bào cơ xương và/hoặc tế bào cơ tim) tăng lên so với khả năng lây nhiễm vào tế bào cơ bởi virion AAV chứa protein capsit AAV mẹ không biến đổi. Theo các khía cạnh liên quan khác của sáng chế, protein capsit AAV biến thể, khi có mặt trong virion AAV cũng tăng cường khả năng chống lại trung hòa bằng kháng thể kháng AAV.

Theo một vài khía cạnh của sáng chế, các virion AAV tái tổ hợp (rAAV) được đề xuất, các virion rAAV này chứa protein capsit biến thể như được mô tả trong bản

mô tả này, trong đó virion rAAV có khả năng lây nhiễm vào một hoặc nhiều tế bào cơ (chẳng hạn tế bào cơ xương và/hoặc tế bào cơ tim) tăng lên so với khả năng lây nhiễm vào tế bào cơ bởi virion AAV chứa protein capsit AAV mẹ không được biến đổi tương ứng. Theo một số phương án, virion rAAV có khả năng lây nhiễm tăng lên ở tất cả các tế bào cơ so với virion AAV chứa protein capsit AAV mẹ. Theo các phương án khác, virion rAAV có khả năng lây nhiễm trong các loại tế bào cơ nhất định tăng lên nhưng tế bào khác thì không so với virion AAV chứa protein capsit AAV mẹ. Nói cách khác, virion rAAV có khả năng lây nhiễm tăng lên ưu tiên cho một số loại tế bào cơ nhất định nhưng các loại tế bào khác thì không, ví dụ, rAAV chứng minh khả năng lây nhiễm tốt hơn vào một hoặc nhiều loại tế bào được chọn từ nguyên bào sợi cơ xương, tế bào vệ tinh cơ xương, nguyên bào sợi tim, tế bào tiền thân tim, tế bào cơ trơn và/hoặc tế bào cơ hoành, nhưng không chứng minh khả năng lây nhiễm ở tất cả các loại tế bào.

Theo một số phương án, virion rAAV chứa axit nucleic khác loại. Theo một số phương án, axit nucleic khác loại mã hóa cho ARN mà mã hóa cho polypeptit. Theo các phương án khác, trình tự axit nucleic khác loại mã hóa cho ARN không mã hóa cho polypeptit, ví dụ, trình tự axit nucleic khác loại là tác nhân can thiệp ARN, ARN hướng dẫn cho nucleaza, v.v..

Cũng được đề xuất trong sáng chế là dược phẩm chứa virion rAAV lây nhiễm và chất mang dược dụng.

Cũng được đề xuất trong bản mô tả này là việc sử dụng virion rAAV chứa protein capsit biến thể như đã mô tả trong bản mô tả này trong phương pháp vận chuyển axit nucleic khác loại đến tế bào đích (chẳng hạn như tế bào cơ tim) bằng cách cho tế bào đích tiếp xúc với virion rAAV. Theo một vài phương án khác, tế bào đích là *in vivo*, chẳng hạn như trong tim của cá thể cần điều trị bệnh tim mạch. Theo phương án khác, tế bào đích là *in vitro*.

Cũng được đề xuất là phương pháp điều trị và/hoặc ngăn ngừa bệnh (ví dụ, rối loạn cơ xương hoặc cơ tim) bằng cách cho đối tượng cần điều trị dùng một lượng có hiệu quả của virion rAAV chứa protein capsit biến thể như đã mô tả trong bản mô tả này hoặc dược phẩm chứa một lượng có hiệu quả điều trị của virion rAAV.

Cũng được đề xuất trong sáng chế là axit nucleic được phân lập chứa trình tự mã hóa cho protein capsit AAV biến thể như đã mô tả trong bản mô tả này và tế bào chủ chứa axit nucleic đã phân lập. Theo phương án khác nữa, axit nucleic được phân lập và/hoặc tế bào chủ đã phân lập chứa rAAV.

Theo một vài khía cạnh, protein capsit AAV biến thể chứa đột biến chèn từ 5 đến khoảng 20 axit amin ("peptit khác loại" hoặc "đoạn chèn peptit") vào trong vòng GH của protein capsit, so với protein capsit AAC mẹ tương ứng, trong đó protein capsit biến thể, khi có mặt trong virion AAV, làm tăng khả năng lây nhiễm vào trong tế bào cơ so với khả năng lây nhiễm vào tế bào cơ bởi virion AAV chứa protein capsit AAV mẹ tương ứng. Theo một số phương án, peptit bao gồm hoặc gần như chỉ bao gồm trình tự được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25),ASDSTKA (SEQ ID NO:26), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32), LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38), LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), và LAASDSTKAA (SEQ ID NO:42). Theo một số phương án được ưu tiên, peptit bao gồm hoặc gần như chỉ bao gồm trình tự được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29) và LATNKIGVTS (SEQ ID NO:37).

Theo một số khía cạnh, protein capsit AAV biến thể chứa một hoặc nhiều đột biến thay thế axit amin so với protein capsit AAV mẹ tương ứng, trong đó protein capsit biến thể, khi có mặt trong virion AAV, làm tăng khả năng lây nhiễm vào trong

tế bào cơ so với khả năng lây nhiễm trong tế bào cơ bởi virion AAV chứa protein capsit AAV mẹ tương ứng.

Theo một số phương án, protein capsit AAV biến thể được mô tả chứa đột biến thay thế P363L so với AAV2 và tùy ý còn có thêm đột biến thay thế E347K và/hoặc V708I so với AAV2.

Theo phương án liên quan, protein capsit AAV chứa việc thêm peptit và một hoặc nhiều đột biến thay thế axit amin so với protein capsit AAV mẹ tương ứng, trong đó protein capsit biến đổi, khi có mặt trong virion AAV, làm tăng khả năng lây nhiễm vào trong tế bào cơ so với khả năng lây nhiễm vào trong tế bào cơ bởi virion AAV chứa protein capsit AAV mẹ tương ứng. Theo một vài phương án, protein capsit AAV biến thể được đề xuất chứa đột biến chèn peptit và đột biến thay thế V708I so với AAV2, trong đó đột biến chèn peptit tùy ý được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25),ASDSTKA (SEQ ID NO:26), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32), LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38), LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), và LAASDSTKAA (SEQ ID NO:42), tốt hơn nếu được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29) và LATNKIGVTS (SEQ ID NO:37). Theo một vài phương án, protein capsit AAV biến thể được đề xuất chứa đột biến chèn peptit và đột biến thay thế P363L so với AAV2, trong đó đột biến chèn peptit được chọn tùy ý từ nhóm chỉ bao gồm GNLTKGN (SEQ ID NO:16), LAGNLTKGNA (SEQ ID NO:30),

QADTTKN (SEQ ID NO:23) và LAQADTTKNA (SEQ ID NO:39).

Theo một vài phương án, protein capsit AAV biến thể được mô tả chứa peptit khác loại LANKIQRTDA (SEQ ID NO:27) và đột biến thay thế V708I so với AAV2 và còn tùy ý chứa thêm đột biến thay thế A593E và/hoặc S109T và/hoặc T330A và/hoặc R588M so với AAV2. Theo một số phương án khác, protein capsit AAV biến thể được mô tả trong bản mô tả này có chứa peptit khác loại (SEQ ID NO:27) và đột biến thay thế A35P so với AAV2. Theo các phương án khác, protein capsit AAV biến thể được mô tả chứa peptit khác loại LANKIQRTDA (SEQ ID NO:27) và đột biến thay thế axit amin N312K, N449D, N551S, I698V, và L735Q so với AAV2 và còn tùy ý có thêm đột biến V708I so với AAV2.

Theo một vài phương án, protein capsit AAV biến thể được mô tả chứa peptit khác loại LANKTTNKDA (SEQ ID NO:28) và đột biến thay thế V708I so với AAV2 và tùy ý còn có thêm đột biến thay thế S109T và/hoặc W694C và/hoặc W606C so với AAV2. Theo các phương án khác, protein capsit AAV biến thể được mô tả chứa peptit khác loại LANKTTNKDA (SEQ ID NO:28) và đột biến thay thế I698V so với AAV2. Theo các phương án khác, protein capsit AAV biến thể được mô tả chứa peptit khác loại LANKTTNKDA (SEQ ID NO:28) và đột biến thay thế axit amin N312K, N449D, N551S, I698V, và L735Q so với AAV2 và còn tùy ý có thêm đột biến V708I so với AAV2.

Theo một vài phương án, protein capsit AAV biến thể được mô tả chứa peptit khác loại LATNKIGVTA (SEQ ID NO:29) và đột biến thay thế V708I so với AAV2 và tùy ý còn có thêm đột biến thay thế N449K và/hoặc G222S so với AAV2. Theo các phương án khác, protein capsit AAV biến thể được mô tả chứa peptit khác loại LANKTTNKDA LATNKIGVTA (SEQ ID NO:29) và đột biến thay thế axit amin N312K, N449D, N551S, I698V, và L735Q so với AAV2 và còn tùy ý có thêm đột biến V708I so với AAV2.

Theo một vài phương án, protein capsit AAV biến thể được mô tả chứa peptit khác loại như được mô tả trong bản mô tả này và đột biến thay thế P363L so với AAV2.

Cũng được đề xuất trong bản mô tả này là phương pháp sản xuất và/hoặc vận

chuyển rAAV có chứa capsit AAV biến thể như được mô tả trong bản mô tả này. Ngoài ra, được đề xuất trong bản mô tả này là kit chứa rAAV chứa capsit AAV biến thể như được mô tả trong bản mô tả này và để sử dụng trong phương pháp được mô tả trong bản mô tả này.

Theo phương án khác, virion AAV chứa protein capsit biến thể trong các đoạn trước đó có thể kết hợp các phương án bất kỳ được đề cập trước đó hoặc sau đó. Thật vậy, đánh giá cao rằng các khía cạnh nhất định của sáng chế, nhằm rõ ràng, được mô tả trong nội dung của các phương án riêng rẽ, cũng có thể được đề xuất ở dạng kết hợp trong một phương án duy nhất. Các dấu hiệu trái ngược, khác nhau theo sáng chế, nhằm ngắn gọn, được mô tả trong nội dung của một phương án duy nhất, cũng có thể được đề xuất riêng rẽ hoặc ở dạng kết hợp phụ bất kỳ phù hợp. Tất cả các dạng kết hợp của các phương án liên quan đến sáng chế đều nằm cụ thể trong phạm vi của sáng chế và được đề cập trong bản mô tả này như thể mỗi dạng kết hợp và mọi dạng kết hợp được bộc lộ một cách riêng rẽ và rõ ràng. Ngoài ra, tất cả các dạng kết hợp phụ của các phương án và yếu tố khác nhau cũng bao gồm cụ thể trong sáng chế và được đề cập trong bản mô tả này như thể mỗi dạng kết hợp và mọi dạng kết hợp phụ được bộc lộ một cách riêng rẽ và rõ ràng.

Phần bản chất kỹ thuật của sáng chế không được dự tính để xác định yêu cầu bảo hộ cũng như không nhằm giới hạn phạm vi của sáng chế theo cách bất kỳ.

Các dấu hiệu và lợi ích khác theo sáng chế được đề cập trong bản mô tả này sẽ trở nên rõ ràng với phần Hình vẽ, Mô tả chi tiết sáng chế, và Yêu cầu bảo hộ.

Mô tả vắn tắt các hình vẽ

Sáng chế được hiểu rõ nhất với phần mô tả chi tiết sau khi đọc kết hợp với các hình vẽ kèm theo. Bằng sáng chế hoặc đơn yêu cầu cấp bằng sáng chế này bao gồm ít nhất một hình vẽ được dựng bằng màu. Các bản sao của bằng sáng chế hoặc đơn yêu cầu cấp bằng sáng chế này với (các) hình vẽ màu sẽ được cung cấp bởi Văn phòng khi có yêu cầu và thanh toán một khoản phí cần thiết. Cần nhấn mạnh rằng, theo ứng dụng thực tế chung, các nét khác nhau của hình vẽ không theo quy mô. Ngược lại, nhằm rõ ràng, kích thước của các nét khác nhau được mở rộng hoặc thu hẹp một cách tùy ý. Bản vẽ bao gồm các hình vẽ sau.

Fig. 1 thể hiện các phương án của phương pháp tiến hóa trực hệ. Bước (a) thể hiện việc tạo ra thư viện capsit virut bao gồm các tổ hợp của kỹ thuật đột biến ADN và gen *cap*. Bước (b) thể hiện việc đóng gói virut sao cho mỗi hạt virut bao gồm capsit đột biến bao quanh gen *cap* mã hóa capsit này và được tinh chế. Sau đó, thư viện capsit được đặt trong điều kiện áp suất chọn lọc *in vitro* hoặc *in vivo*. Theo khía cạnh này về công nghệ tiến hóa trực hệ, mô hoặc vật liệu tế bào quan tâm được thu hoạch để phân lập biến thể AAV mà đã lây nhiễm thành công đích này, và các virut thành công được thu hồi. Bước (c) thể hiện sự làm giàu Giai đoạn 1 các dòng vô tính thành công qua việc chọn lọc lặp lại. Bước (d) thể hiện sự làm giàu giai đoạn 2 gen *cap* đã chọn mà tham gia các bước đa dạng hóa lại và chọn lọc tiếp để làm tăng tiếp sức mạnh của virut. Bước (e) thể hiện các biến thể, được nhận diện dưới dạng các lần đột biến (Hits) trong quá trình các giai đoạn chọn lọc vectơ 1 và 2, mà sẽ được tạo ra dưới dạng vectơ AAV tái tổ hợp và xác định đặc điểm mức tải nạp của loại tế bào và các đích mô khác nhau. Bằng quy trình tiến hóa trực hệ AAV tự nhiên, đã chứng minh được các biến thể được đề cập trong bản mô tả này có khả năng tải nạp tế bào cơ và vận chuyển hệ gen (hệ gen mã hóa gen *cap* biến thể) trong suốt quy trình chọn lọc.

Fig. 2 trình bày kết quả khuếch đại PCR hệ gen virut từ mô cơ xương và mô cơ tim từ vòng chọn lọc đại diện. Các băng trong hộp đỏ đại diện cho việc khuếch đại hệ gen virut thành công.

Fig. 3A-3C trình bày tần suất của các môtip trong phân tích trình tự. Fig. 3A thể hiện phân tích trình tự vòng 4 dùng cho áp lực chọn lọc của vận chuyển trong tĩnh mạch đến cơ tim. Fig. 3B thể hiện phân tích trình tự vòng 2 dùng cho áp lực chọn lọc của vận chuyển trong tĩnh mạch với sự có mặt của các kháng thể trung hòa đến cơ tim. Fig. 3C thể hiện phân tích trình tự Vòng 3 dùng cho áp lực chọn lọc của vận chuyển trong tĩnh mạch đến cơ xương. Fig 3A cho thấy 57,40% là Motif LANKIQRTDA, 16,96% là Motif LANKTTNKDA, 7,32% Motif A593E, 7,32% những cái khác, 4,88% Motif V708I và 4,88% Motif LASNTVKAIA. Fig 3B cho thấy 21,14% những cái khác, 20,33% Motif LAQADTTKNA, 15,45% Motif LANKTTNKDA, 15,45% Motif LAASNITKAA, 15,45% Motif thể khảm AAV6/AAV5 và 12,20% Motif LANTVKLSTA. Fig 3C cho thấy 43,21% Motif A593E, 41,98% Motif P363L và 14,81% những cái khác.

Fig. 4A-4C là mô hình ba chiều đại diện của AAV2 chứa heptame ngẫu nhiên sau axit amin 587 và đột biến thay thế V708I. Fig 4B là mô hình ba chiều đại diện của thể khảm AAV6/AAV5 chứa đột biến thay thế V229I, A490T, và A581T (tương ứng với trình tự axit amin nêu trong SEQ ID NO:62). Fig. 4C là mô hình ba chiều đại diện của AAV2 chứa đột biến thay thế P363L.

Fig. 5 thể hiện sự sắp thẳng hàng của các SEQ ID NO:1-11 AAV kiểu dại có các vị trí axit amin nằm giữa và nằm ngang các kiểu huyết thanh AAV1, AAV2, AAV3A, AAV3B, và AAV4-10 kiểu dại (xuất hiện trong tự nhiên).

Fig. 6A-6E cung cấp dữ liệu về việc tải nạp tế bào cơ tim người *in vitro* bằng virut AAV tái tổ hợp chứa capsit biến thể AAV mới LANKIQRTDA+V708I, capsit biến thể AAV mới LANKTTNKDA+V708I, và capsit biến thể LATNKIGVTA+V708I mới, mỗi cái biểu hiện gen chuyển GFP dưới sự kiểm soát của vùng gen khởi đầu CAG. Fig. 6A: các tế bào được biệt hóa thành tế bào cơ tim từ dòng tế bào gốc toàn năng người được cho nhiễm vào trong biến thể AAV mới LANKIQRTDA+V708I.CAG.GFP, biến thể AAV mới LANKTTNKDA+V708I.CAG.GFP, biến thể AAV mới LATNKIGVTA+V708I.CAG.GFP hoặc đối chứng là kiểu dại AAV1.CAG.GFP, AAV2.CAG.GFP, và AAV9.CAG.GFP tại các MOI là 20, 100, 500, và 2500. Hình ảnh huỳnh quang của môi trường nuôi cấy tế bào sau 6 ngày lây nhiễm tại tất cả các MOI chứng tỏ rằng biến thể AAV mới tải nạp tế bào cơ tim tốt hơn so với capsit AAV1, AAV2, hoặc AAV9 kiểu dại. Fig. 6B: định lượng phần trăm tế bào cơ tim dương tính GFP trong mỗi môi trường nuôi cấy bằng đếm dòng chảy tế bào chứng tỏ rằng biến thể AAV mới mang lại sự cải thiện đáng kể, độc lập với liều lượng về số lượng tế bào được tải nạp so với capsit AAV1, AAV2, hoặc AAV9 kiểu dại. * $p < 0.05$. Fig. 6C-6D định lượng hàm lượng GFP trong mỗi môi trường nuôi cấy bằng Western blot chứng tỏ rằng biến thể AAV mời mang lại sự cải thiện đáng kể trong biểu hiện của gen chuyển so với capsit AAV1, AAV2, hoặc AAV9 kiểu dại. NT= không được chuyển nap. Fig. 6E: các tế bào được biệt hóa thành tế bào cơ tim từ dòng tế bào gốc toàn năng người được lây nhiễm với biến thể AAV mới LANKIQRTDA+V708I.CAG.GFP, biến thể AAV mới LANKTTNKDA+V708I.CAG.GFP, biến thể AAV mới

LATNKIGVTA+V708I.CAG.GFP hoặc đối chứng là kiểu dại AAV1.CAG.GFP, AAV2.CAG.GFP, và AAV9.CAG.GFP. Hình ảnh huỳnh quang của môi trường nuôi cấy tế bào sau vào 1, 2, 3, và 5 ngày sau khi lây nhiễm tại MOI 500 chứng tỏ rằng biến thể AAV mới tải nạp tế bào cơ tim tốt hơn và bắt đầu biểu hiện gen chuyển GFP sớm hơn so với capsit AAV1, AAV2, hoặc AAV9 kiểu dại.

Fig. 7A-E cung cấp dữ liều về việc tải nạp tế bào cơ tim người *in vitro* bằng virut AAV tái tổ hợp chứa biến thể AAV mới là capsit thể khảm AAV6/AAV5 có trình tự nêu trong SEQ ID NO: 62, biểu hiện gen chuyển GFP dưới sự kiểm soát của vùng gen khởi đầu CAG. Fig. 7A: các tế bào được biệt hóa thành tế bào cơ tim từ dòng tế bào gốc toàn năng người được cho nhiễm vào capsit thể khảm AAV6/AAV5 biến thể AAV mới hoặc đối chứng là kiểu dại AAV1.CAG.GFP, AAV8.CAG.GFP, và AAV9.CAG.GFP tại MOIs là 20, 100, 500, và 2500. Hình ảnh huỳnh quang của môi trường nuôi cấy tế bào sau 6 ngày lây nhiễm tại tất cả các MOI chứng tỏ rằng biến thể AAV mới tải nạp tế bào cơ tim tốt hơn so với capsit AAV1, AAV8, hoặc AAV9 kiểu dại. Fig. 7B: định lượng phần trăm tế bào cơ tim dương tính GFP trong mỗi môi trường nuôi cấy bằng đếm dòng chảy tế bào chứng tỏ rằng biến thể AAV mới mang lại sự cải thiện đáng kể, độc lập với liều lượng về số lượng tế bào được tải nạp so với capsit AAV1, AAV8, hoặc AAV9 kiểu dại. * $p < 0.05$. Fig. 7C-7D: định lượng hàm lượng GFP trong mỗi môi trường nuôi cấy bằng Western blot chứng tỏ rằng biến thể AAV mời mang lại sự cải thiện đáng kể trong biểu hiện của gen chuyển so với capsit AAV1, AAV8, hoặc AAV9 kiểu dại. Tá dược lỏng = không được tải nạp. Fig. 7E: các tế bào được biệt hóa thành tế bào cơ tim từ dòng tế bào gốc toàn năng người được lây nhiễm vào capsit thể khảm AAV6/AAV5 biến thể AAV mới hoặc đối chứng là kiểu dại AAV8.CAG.GFP. Hình ảnh huỳnh quang của môi trường nuôi cấy tế bào sau vào 3, 4, 5 và 6 ngày sau khi lây nhiễm tại MOI 2500 chứng tỏ rằng biến thể AAV mới tải nạp tế bào cơ tim tốt hơn và bắt đầu biểu hiện gen chuyển GFP sớm hơn so với capsit AAV8 kiểu dại.

Fig. 8A-C cung cấp dữ liều về việc tải nạp tơ cơ xương người *in vitro* bằng virut AAV tái tổ hợp chứa biến thể AAV mới LANKIQRTDA+V708I capsit, biến thể AAV mới LANKTTNKDA+V708I capsit, và capsit thể khảm AAV6/AAV5 biến thể AAV mới, mỗi cái biểu hiện gen chuyển GFP dưới sự kiểm soát của vùng gen khởi

đầu CAG. Fig. 8A: các tế bào được biệt hóa thành tơ cơ xương người từ dòng tế bào gốc toàn năng người được cho nhiễm vào trong biến thể AAV mới LANKIQRTDA+V708I.CAG.GFP, biến thể AAV mới LANKTTNKDA+V708I.CAG.GFP, biến thể AAV mới AAV6/AAV5 thể khảm.CAG.GFP hoặc đối chứng là kiểu dại AAV8.CAG.GFP và AAV9.CAG.GFP tại MOI là 100, 500, và 2500. Hình ảnh huỳnh quang của môi trường nuôi cấy tế bào sau 7 ngày lây nhiễm tại tất cả các MOI chứng tỏ rằng biến thể AAV mới tải nạp tơ cơ xương tốt hơn so với capsit AAV8 hoặc AAV9 kiểu dại. Fig. 8B: định lượng phần trăm tơ cơ xương dương tính GFP trong mỗi môi trường nuôi cấy bằng đếm dòng chảy tế bào chứng tỏ rằng biến thể AAV mới mang lại sự cải thiện đáng kể, độc lập với liều lượng về số lượng tế bào được tải nạp so với capsit AAV8 hoặc AAV9 kiểu dại. * $p < 0.05$. Fig. 8C: các tế bào được biệt hóa thành tơ cơ xương người từ dòng tế bào gốc toàn năng người được lây nhiễm với biến thể AAV mới LANKIQRTDA+V708I.CAG.GFP, biến thể AAV mới AAV6/AAV5 thể khảm.CAG.GFP hoặc đối chứng là kiểu dại AAV8.CAG.GFP và AAV9.CAG.GFP. Hình ảnh huỳnh quang của môi trường nuôi cấy tế bào sau vào 2-7 ngày sau khi lây nhiễm tại MOI 2500 chứng tỏ rằng biến thể AAV mới tải nạp tơ cơ xương tốt hơn và bắt đầu biểu hiện gen chuyển GFP sớm hơn so với capsit AAV8 hoặc AAV9 kiểu dại.

Fig. 9A-B cung cấp dữ liệu về việc tải nạp tế bào đầu dòng cơ người *in vitro* bằng virut AAV tái tổ hợp chứa biến thể AAV mới LANKIQRTDA+V708I capsit, biến thể AAV mới LANKTTNKDA+V708I capsit, và capsit thể khảm AAV6/AAV5 biến thể AAV mới, mỗi cái biểu hiện gen chuyển GFP dưới sự kiểm soát của vùng gen khởi đầu CAG. Fig. 9A: các tế bào được biệt hóa thành tế bào đầu dòng cơ từ dòng tế bào gốc toàn năng người được cho nhiễm vào trong biến thể AAV mới LANKIQRTDA+V708I.CAG.GFP, biến thể AAV mới LANKTTNKDA+V708I.CAG.GFP, biến thể AAV mới AAV6/AAV5 thể khảm.CAG.GFP hoặc đối chứng là kiểu dại AAV9.CAG.GFP tại MOI là 500. Hình ảnh huỳnh quang của môi trường nuôi cấy tế bào sau 6 ngày lây nhiễm tại tất cả các MOI chứng tỏ rằng biến thể AAV mới tải nạp tế bào đầu dòng cơ tốt hơn so với capsit AAV9 kiểu dại. Fig. 9B: định lượng phần trăm tế bào đầu dòng cơ dương tính GFP trong mỗi môi trường nuôi cấy bằng đếm dòng chảy tế bào chứng tỏ rằng biến thể

AAV mới mang lại sự cải thiện đáng kể, độc lập với liều lượng về số lượng tế bào được tải nạp so với capsit AAV9 kiểu dại. * $p < 0.05$.

Fig. 10A-B cung cấp dữ liều về việc tải nạp tế bào đầu dòng cơ người *in vitro* bằng virut AAV tái tổ hợp chứa biến thể AAV mới LANKIQRTDA+V708I capsit, biến thể AAV mới LANKTTNKDA+V708I capsit, và capsit thể khảm AAV6/AAV5 biến thể AAV mới, mỗi cái biểu hiện gen chuyển GFP dưới sự kiểm soát của vùng gen khởi đầu CAG. Fig. 10A: Tăng gấp nhiều lần trong việc tải nạp tế bào cơ tim người bởi các biến thể capsit AAV mới so với AAV8 và AAV9 kiểu dại, các kiểu huyết thanh được sử dụng rộng rãi nhất trong ứng dụng lâm sàng đối với các bệnh cơ. Fig. 10B: Tăng nhiều lần trong tải nạp tơ cơ xương người bởi biến thể capsit AAV mới so với AAV8 và AAV9 kiểu dại.

Fig. 11A-B cung cấp dữ liệu về việc tải nạp ở mô chuột *in vivo* bằng virut AAV tái tổ hợp chứa biến thể AAV mới capsit LANKIQRTDA+V708I biểu hiện gen chuyển luciferaza dưới sự điều khiển của vùng gen khởi đầu CAG. Chuột được tiêm một mũi tiêm tĩnh mạch thông qua tĩnh mạch đuôi $2x10^{11}$ hệ gen virut trên mỗi con. Fig.11A: Hình ảnh của luciferaza tại 14 ngày (bên trái) và 28 ngày (bên phải) sau khi cung cấp chứng tỏ rằng biến thể AAV mới capsit LANKIQRTDA+V708I có thể tải nạp tế bào chuột *in vivo*. Fig. 11B: phạm vi hoạt động của luciferaza trong tim, cơ hoành, và cơ bốn đầu sau 56 ngày cung cấp chứng tỏ rằng biến thể AAV mới capsit LANKIQRTDA+V708I có thể tải nạp tế bào cơ xương và cơ tim của chuột *in vivo*.

Fig. 12A-B cung cấp dữ liệu tải nạp cơ xương ở linh trưởng không phải con người *in vivo* bằng virut AAV tái tổ hợp chứa biến thể AAV mới capsit LANKIQRTDA+V708I biểu hiện gen chuyển GFP dưới sự điều khiển của vùng gen khởi đầu CAG. Linh trưởng không phải con người được tiêm 3 mũi tiêm tĩnh mạch thông qua tĩnh mạch đuôi 10^{11} hệ gen virut vào mỗi cơ đùi ngoài bên trái, và mô cơ được phân tích sau 4 tuần cung cấp. Fig. 12A: hình ảnh của haemotoxylin và eosin (H&E) và nhuộm màu kháng thể kháng GFP của mặt cắt ngang vị trí sinh thiết gần nhất tại độ phóng đại 2x, 4x, và 20x chứng tỏ rằng biến thể AAV mới capsit LANKIQRTDA+V708I có thể tải nạp tế bào cơ xương ở linh trưởng *in vivo*. Fig. 12B: hình ảnh của haemotoxylin và eosin (H&E) và nhuộm màu kháng thể kháng GFP của mặt cắt dọc vị trí sinh thiết gần nhất tại độ phóng đại 2x, 4x, và 20x chứng tỏ rằng biến

thể AAV mới capsit LANKIQRTDA+V708I có thể tải nạp tế bào cơ xương ở linh trưởng *in vivo*.

Mô tả chi tiết sáng chế

Trước khi phương pháp và dược phẩm được mô tả, cần phải hiểu rằng sáng chế này không bị giới hạn ở phương pháp hoặc dược phẩm cụ thể được mô tả và do đó có thể thay đổi được. Cũng cần hiểu rằng các thuật ngữ được sử dụng trong bản mô tả này chỉ nhằm mục đích mô tả các phương án cụ thể, và không dự định giới hạn phạm vi của sáng chế, do phạm vi của sáng chế sẽ chỉ bị giới hạn ở các điểm yêu cầu bảo hộ kèm theo.

Sáng chế được đề cập trong bản mô tả này được minh họa trong các hình vẽ và phần mô tả. Mặc dù các phương án cụ thể được minh họa trong các hình vẽ, tuy nhiên sáng chế không dự định giới hạn ở các phương án hoặc phương án cụ thể được minh họa và/hoặc đề cập. Đúng hơn, sáng chế được đề cập trong bản mô tả này dự định bao gồm tất cả các biến đổi, các cấu trúc thay thế, và dạng tương đương đều nằm trong phạm vi của sáng chế. Như vậy, các hình vẽ được dự định nhằm minh họa và không giới hạn phạm vi của sáng chế.

Khi các khoảng giá trị được đưa ra, cần hiểu rằng mỗi giá trị xen giữa, tính đến phần mười của đơn vị giới hạn dưới, trừ khi có quy định rõ ràng khác trong bản mô tả, giữa giới hạn trên và giới hạn dưới của khoảng đó cũng được đề cập cụ thể. Mỗi khoảng nhỏ hơn giữa các giá trị hoặc giá trị xen giữa đã nêu bất kỳ trong khoảng đã nêu và giá trị đã nêu hoặc xen giữa khác bất kỳ trong khoảng đã nêu đó cũng nằm trong phạm vi của sáng chế. Giới hạn trên và giới hạn dưới của các khoảng nhỏ hơn này có thể độc lập nằm trong hoặc không nằm trong khoảng này, và mỗi khoảng có một trong hai giới hạn này, không có cả hai giới hạn này hoặc cả hai giới hạn này chứa các khoảng nhỏ nhất cũng nằm trong phạm vi của sáng chế, đối tượng cho giới hạn được loại trừ cụ thể bất kỳ trong khoảng đã nêu. Nếu khoảng đã nêu bao gồm một hoặc cả hai giới hạn, các khoảng không bao gồm một trong hai giới hạn hoặc cả hai giới hạn này cũng nằm trong phạm vi của sáng chế.

Trừ khi có định nghĩa khác, tất cả các thuật ngữ kỹ thuật và khoa học được sử dụng trong bản mô tả này đều có nghĩa giống như được hiểu thông thường bởi người

có hiểu biết trung bình trong lĩnh vực kỹ thuật mà sáng chế có liên quan. Mặc dù các phương pháp và vật liệu bất kỳ tương tự hoặc tương đương với các phương pháp và vật liệu được đề cập trong bản mô tả này có thể được sử dụng trong ứng dụng thực tế hoặc thử nghiệm sáng chế, một số phương pháp và vật liệu tiềm năng và ưu tiên được mô tả dưới đây. Tất cả các công bố được đề cập trong bản mô tả này được đưa vào bản mô tả này bằng cách viện dẫn để bộc lộ và mô tả các phương pháp và/hoặc vật liệu liên quan đến những gì mà các công bố đã trích dẫn. Cần hiểu rằng sáng chế không sử dụng phần bộc lộ bất kỳ của công bố đã được đưa vào trong trường hợp có mâu thuẫn.

Người có hiểu biết trong lĩnh vực kỹ thuật này khi đọc bản mô tả này sẽ hiển nhiên hiểu rằng mỗi phương án riêng rẽ được mô tả và minh họa trong bản mô tả này có thành phần và dấu hiệu riêng biệt có thể được tách rời hoặc kết hợp một cách dễ dàng với các dấu hiệu của các phương án bất kỳ trong số một vài phương án khác mà không tách rời khỏi phạm vi của sáng chế. Các phương pháp bất kỳ được nêu có thể được thực hiện theo trình tự trường hợp được nêu hoặc theo trình tự khác bất kỳ có thể xảy ra hợp lý.

Lưu ý rằng như được sử dụng trong bản mô tả này và trong yêu cầu bảo hộ đi kèm, các dạng số ít "a," "an," và "the" bao gồm các tham chiếu số nhiều trừ khi ngữ cảnh chỉ rõ khác. Do đó, ví dụ khi nói "virion AAV tái tổ hợp" bao gồm dạng số nhiều của các hạt virut này và nói đến "tế bào cơ" bao gồm nói đến một hoặc nhiều tế bào cơ hoặc cái tương đương mà những người có hiểu biết trung bình trong lĩnh vực kỹ thuật biết đến, và v.v.. Cũng cần lưu ý rằng yêu cầu bảo hộ có thể được soạn thảo để loại trừ yếu tố tùy ý bất kỳ. Như vậy, sự thể hiện này được dự định để dùng như cơ sở tiền lệ nhằm sử dụng thuật ngữ độc quyền như "duy nhất", "chỉ" và các thuật ngữ tương tự liên quan đến việc mô tả các yếu tố yêu cầu bảo hộ, hoặc sử dụng giới hạn "phủ định".

Các công bố được đề cập trong bản mô tả này chỉ được cung cấp cho phần mô tả của chúng trước ngày nộp đơn của đơn sáng chế. Cần hiểu rằng tất cả phần bộc lộ trong bản mô tả này không phải là sự thừa nhận sáng chế không được quyền xuất hiện trước tài liệu công bố này theo hiệu lực của giải pháp kỹ thuật đã biết. Ngoài ra, các ngày công bố được đưa ra có thể khác với các ngày công bố thực tế có thể cần được xác nhận độc lập.

Định nghĩa

Virut adeno kết hợp là parovirut không gây bệnh bao gồm hệ gen ADN sợi đơn 4,7 kb nằm trong capsit hai mươi mặt, không được bọc. Hệ gen chứa ba khung đọc mở (ORF) nằm bên sườn đầu lặp lặp lại đảo ngược (ITR) đóng vai trò như là điểm khởi đầu sao chép của virut và tín hiệu đóng gói. ORF *rep* mã hóa cho bốn protein phi cấu trúc đóng vai trò trong sao chép virut, điều hòa phiên mã, tích hợp đặc hiệu vị trí và lắp ráp hạt virut. ORF *cap* mã hóa cho ba protein cấu trúc (VP1-3) lắp ráp để tạo thành capsit virut 60-mer. Cuối cùng, ORF xuất hiện dưới dạng khung đọc xen kẽ trong gen *cap* sản xuất protein kích hoạt lắp ráp (AAP), protein virut xác định vị trí protein capsit AAV trong nhân và thực hiện chức năng trong quá trình lắp ráp capsit.

Có một vài kiểu huyết thanh có trong trong tự nhiên ("kiểu dại") và trên 100 biến thể đã biết của AAV, mỗi loại trong số đó khác nhau về trình tự axit amin, cụ thể là tại vùng siêu biến của protein capsit, và do đó khác nhau về đặc tính vận chuyển gen. Không có AAV nào liên quan đến bệnh ở người, khiến cho AAV tái tổ hợp trở nên thu hút đối với các ứng dụng lâm sàng.

Với mục đích của sáng chế, thuật ngữ "AAV" là viết tắt của virut adeno kết hợp, bao gồm, nhưng không chỉ giới hạn ở, chính bản thân virut và các dẫn xuất của nó. Trừ khi được chỉ ra khác, thuật ngữ này dùng để chỉ tất cả tập con hoặc kiểu huyết thanh và cả dạng tái tổ hợp và có khả năng sao chép. Thuật ngữ "AAV" bao gồm, nhưng không chỉ giới hạn ở, AAV kiểu 1 (AAV-1 hoặc AAV1), AAV kiểu 2 (AAV-2 hoặc AAV2), AAV kiểu 3A (AAV-3A hoặc AAV3A), AAV kiểu 3B (AAV-3B hoặc AAV3B), AAV kiểu 4 (AAV-4 hoặc AAV4), AAV kiểu 5 (AAV-5 hoặc AAV5), AAV kiểu 6 (AAV-6 hoặc AAV6), AAV kiểu 7 (AAV-7 hoặc AAV7), AAV kiểu 8 (AAV-8 hoặc AAV8), AAV kiểu 9 (AAV-9 hoặc AAV9), AAV kiểu 10 (AAV-10 hoặc AAV10 hoặc AAVrh10), AAV chim, AAV bò, AAV chó, AAV dê, AAV ngựa, AAV linh trưởng, AAV không phải linh trưởng, và AAV cừu. "AAV linh trưởng" dùng để chỉ AAV mà lây nhiễm linh trưởng, "AAV không phải linh trưởng" dùng để chỉ AAV mà lây nhiễm cho động vật có vú không phải linh trưởng, "AAV bò" dùng để chỉ AAV lây nhiễm cho bò, v.v..

Trình tự hệ gen của các kiểu huyết thanh khác nhau của AAV, cũng như trình tự của trình tự các đầu lặp lại (TR) tự nhiên, protein Rep, và tiểu phần capsit đã được biết đến trong lĩnh vực kỹ thuật. Trình tự này có thể được tìm thấy trong tài liệu hoặc cơ sở

dữ liệu công khai chẳng hạn như GenBank, xem, ví dụ, Số truy cập GenBank NC_002077.1 (AAV1), AF063497.1 (AAV1), NC_001401.2 (AAV2), AF043303.1 (AAV2), J01901.1 (AAV2), U48704.1 (AAV3A), NC_001729.1 (AAV3A), AF028705.1 (AAV3B), NC_001829.1 (AAV4), U89790.1 (AAV4), NC_006152.1 (AA5), AF085716.1 (AAV-5), AF028704.1 (AAV6), NC_006260.1 (AAV7), AF513851.1 (AAV7), AF513852.1 (AAV8) NC_006261.1 (AAV-8), AY530579.1 (AAV9), AAT46337 (AAV10) và AAO88208 (AAVrh10); nội dung của các tài liệu này được đưa vào bản mô tả bằng cách viện dẫn để tham khảo về các trình tự axit nucleic và axit amin AAV. Cũng xem, ví dụ, Srivistava et al. (1983) J. Virology 45:555; Chiorini et al. (1998) J. Virology 71:6823; Chiorini et al. (1999) J. Virology 73: 1309; Bantel-Schaal et al. (1999) J. Virology 73:939; Xiao et al. (1999) J. Virology 73:3994; Muramatsu et al. (1996) Virology 221:208; Shade et. al. (1986) J. Virol. 58:921; Gao et al. (2002) Proc. Nat. Acad. Sci. USA 99: 11854; Moris et al. (2004) Virology 33:375-383; công bố sáng chế quốc tế số WO 00/28061, WO 99/61601, WO 98/11244; và bằng sáng chế Mỹ số 6,156,303.

Trình tự của protein cap (capsit) tồn tại trong tự nhiên liên quan đến kiểu huyết thanh AAV đã được biết đến trong lĩnh vực kỹ thuật và bao gồm các trình tự được mô tả trong bản mô tả này như AAV1 (SEQ ID NO:1), AAV2 (SEQ ID NO:2), AAV3A (SEQ ID NO:3), AAV3B (SEQ ID NO:4), AAV4 (SEQ ID NO:5), AAV5 (SEQ ID NO:6), AAV6 (SEQ ID NO:7), AAV7 (SEQ ID NO:8), AAV8 (SEQ ID NO:9), AAV9 (SEQ ID NO:10), AAV10 (SEQ ID NO:11), và AAVrh10 (SEQ ID NO:12). Thuật ngữ "protein capsit AAV biến thể" hoặc " biến thể AAV" dùng để chỉ protein capsit AAV chứa trình tự axit amin có ít nhất một biến đổi hoặc sự thay thế (bao gồm mất, thêm, đột biến điểm, v.v.) so với trình tự protein capsit AAV "kiểu dại" hoặc tồn tại trong tự nhiên, ví dụ, như trình tự nêu trong các SEQ ID NO:l-12. Protein capsit AAV biến thể có thể tương đồng khoảng 80% hoặc nhiều hơn so với trình tự axit amin của protein capsit kiểu dại, ví dụ, tương đồng 85% hoặc nhiều hơn, tương đồng 90% hoặc nhiều hơn, hoặc tương đồng 95% hoặc nhiều hơn so với trình tự axit amin của protein capsit kiểu dại, ví dụ, tương đồng 98% hoặc 99% so với protein capsit kiểu dại. Protein capsit AAV biến thể có thể không phải là protein capsit kiểu dại.

Với mục đích của sáng chế, "virion AAV" hoặc "hạt virut AAV" dùng để chỉ

hạt virut có chứa ít nhất một protein capsit AAV và polynucleotit AAV được đóng gói.

Với mục đích của sáng chế, thuật ngữ "rAAV" là từ viết tắt dùng để chỉ virut adeno kết hợp tái tổ hợp. "Tái tổ hợp", khi sử dụng cho polynucleottit có nghĩa là polynucleotit là sản phẩm của tổ hợp khác nhau gồm bước tách dòng, cắt giới hạn hoặc nối, và các quá trình khác dẫn đến tạo ra cấu trúc khác so với polynucleotit tìm thấy trong tự nhiên. Virut tái tổ hợp là hạt virut chứa polynucleotit tái tổ hợp. Thuật ngữ tương ứng bao gồm bản sao của cấu trúc polynucleotit ban đầu và con cháu của cấu trúc virut ban đầu.

Thuật ngữ "vectơ rAAV" bao gồm virion rAAV (tức là, hạt virut rAAV) (ví dụ, virion rAAV lây nhiễm), theo định nghĩa bao gồm polynucleotit rAAV; và cũng bao gồm polynucleotit mã hóa cho rAAV (ví dụ, polynucleotit sợi đơn mã hóa rAAV (ss-rAAV); polynucleotit sợi đôi mã hóa rAAV (ds- rAAV), ví dụ, plasmit mã hóa rAAV; và các dạng polynucleotit tương tự).

Nếu virion AAV bao gồm polynucleotit khác loại (tức là polynucleotit khác ngoài hệ gen AAV kiểu dại, ví dụ, gen chuyển được vận chuyển đến tế bào đích, tác nhân ARNi hoặc tác nhân CRISPR để được vận chuyển đến tế bào đích, v..v.), nó thường được gọi là "virion AAV tái tổ hợp (rAAV)" hoặc "hạt virut rAAV". Nói chung, polynucleotit khác loại được kẹp bởi ít nhất một, và thường là bởi hai, đầu lặp lại đảo ngược AAV (ITR).

Thuật ngữ "đóng gói" dùng để chỉ một loạt các sự kiện bên trong tế bào dẫn đến việc lắp ráp và tạo vỏ capsit cho hạt AAV. Các gen "*rep*" và "*cap*" AAV dùng để chỉ trình tự polynucleotit mã hóa protein sao chép và tạo vỏ capsit cho virut adeno kết hợp. Các gen *rep* và *cap* AAV được dùng trong bản mô tả này để chỉ "gen đóng gói" AAV.

Thuật ngữ "virut trợ giúp" cho AAV dùng để chỉ virut cho phép AAV (ví dụ, AAV kiểu dại) được sao chép và đóng gói bởi tế bào động vật có vú. Các dạng khác nhau của virut trợ giúp cho AAV này đã biết trong lĩnh vực kỹ thuật này, bao gồm virut ađeno, virut herpes và poxvirut chẳng hạn như bệnh đậu mùa. Virut ađeno bao gồm một số phân nhóm khác nhau, mặc dù virut ađeno typ 5 của phân nhóm C được sử dụng phổ biến nhất. Rất nhiều virut ađeno có nguồn gốc từ người, động vật có vú

không phải là người và chim đã được biết đến và có sẵn trong các bộ sưu tập được lưu giữ như ATCC. Virut thuộc họ Herpes bao gồm, ví dụ, virut gây bệnh mụn rộp (HSV) và virut Epstein-Barr (EBV), cũng như virut cự bào (CMV) và virut gây bệnh dại giả (PRV); cũng có thể có sẵn trong các bộ sưu tập được lưu giữ như ATCC.

Thuật ngữ "chức năng virut trợ giúp" dùng để chỉ chức năng được mã hóa trong hệ gen virut trợ giúp cho phép AAV sao chép và đóng gói (kết hợp với các yêu cầu khác để sao chép và đóng gói được đề cập trong bản mô tả này). Như được đề cập trong bản mô tả này, "chức năng virut trợ giúp" có thể tạo ra theo một số cách, bao gồm tạo ra virut trợ giúp hoặc tạo ra, ví dụ, trình tự polynucleotit mã hóa cho chức năng cần thiết cho tế bào sản xuất từ xa. Ví dụ, plasmit hoặc vectơ biểu hiện khác bao gồm trình tự nucleotit mã hóa một hoặc nhiều virut ađenoprotein được chuyển nhiễm vào tế bào sản xuất cùng với vectơ rAAV.

Thuật ngữ virut hoặc hạt virut "lây nhiễm" là thuật ngữ bao gồm capsit virut được lắp ráp thành thạo và có khả năng vận chuyển thành phần polynucleotit vào trong tế bào mà các loài virut tính hướng kích thích. Thuật ngữ không nhất thiết hàm ý khả năng sao chép bất kỳ của virut. Thử nghiệm để đếm hạt virut lây nhiễm được mô tả ở nơi nào đó trong bản mô tả này và trong lĩnh vực kỹ thuật. Tính lây nhiễm virut có thể được biểu thị dưới dạng tỷ lệ giữa hạt virut lây nhiễm và tổng hạt virut. Phương pháp xác định tỷ lệ giữa hạt virut lây nhiễm và tổng hạt virut đã được biết đến trong lĩnh vực kỹ thuật. Xem, ví dụ, Grainger et al. (2005) Mol. Ther. 11: S337 (mô tả thử nghiệm chuẩn độ lây nhiễm TCID50); và Zolotukhin et al. (1999) Gene Ther. 6:973. Cũng xem trong phần ví dụ.

Thuật ngữ "tính hướng" như được sử dụng trong bản mô tả này dùng để chỉ việc hướng đích ưu tiên bởi virut (ví dụ, AAV) trong tế bào của các loài vật chủ cụ thể hoặc trong loại tế bào cụ thể của một loài vật chủ. Ví dụ, virut có thể lây nhiễm tế bào của tim, phổi, gan, và cơ có tính hướng rộng hơn (tức là, tăng lên) so với virut chỉ có thể lây nhiễm tế bào phổi và cơ. Tính hướng cũng có thể bao gồm sự phụ thuộc của virut vào loại phân tử bề mặt tế bào cụ thể của vật chủ. Ví dụ, một số virut chỉ có thể lây nhiễm tế bào có glycosaminoglycan bề mặt, trong khi các virut khác chỉ có thể lây nhiễm tế bào có axit sialic (sự phụ thuộc này có thể được kiểm tra bằng cách sử dụng các dòng tế bào khác nhau thiếu hụt các lớp phân tử cụ thể làm tế bào chủ tiềm năng

để lây nhiễm virut). Trong một số trường hợp, tính hướng của virut mô tả sự ưu tiên tương đối của virut. Ví dụ, virut thứ nhất có thể có khả năng lây nhiễm tất cả các loại tế bào nhưng thành công hơn khi lây nhiễm các tế bào có glycosaminoglycan bề mặt. Virut thứ hai có thể được xem là có tính hướng tương tự (hoặc giống hệt) như virut thứ nhất nếu virut thứ hai cũng ưu tiên hơn các đặc điểm tương tự (ví dụ, virut thứ hai cũng thành công hơn khi lây nhiễm các tế bào có glycosaminoglycan bề mặt), ngay cả khi khả năng tải nạp tuyệt đối là không giống nhau. Ví dụ, virut thứ hai có thể có hiệu quả hơn so với virut thứ nhất khi lây nhiễm mọi loại tế bào đã cho được thử nghiệm, nhưng nếu mức ưu tiên tương đối là tương tự (hoặc giống hệt), virut thứ hai cũng có thể được xem là có tính hướng tương tự (hoặc giống hệt) như virut thứ nhất. Theo một số phương án, tính hướng của virion chứa protein capsit AAV biến thể đối tượng không bị thay đổi so với virion có trong tự nhiên. Theo một số phương án, tính hướng của virion chứa protein capsit AAV biến thể đối tượng được kéo dài (tức là, mở rộng) so với virion có trong tự nhiên. Theo một số phương án, tính hướng của virion chứa protein capsit AAV biến thể đối tượng giảm đi so với virion có trong tự nhiên.

Thuật ngữ virut "có khả năng sao chép" (ví dụ, AAV có khả năng sao chép) dùng để chỉ virut lây nhiễm kiểu dại về kiểu hình, và cũng có khả năng được sao chép trong tế bào đã lây nhiễm (tức là với sự có mặt của virut trợ giúp hoặc chức năng virut trợ giúp). Trong trường hợp của AAV, khả năng sao chép thường đòi hỏi sự có mặt của gen đóng gói AAV chức năng. Nói chung, vectơ rAAV như được đề cập trong bản mô tả này không có khả năng sao chép trong tế bào động vật có vú (đặc biệt là ở tế bào người) do thiếu một hoặc nhiều gen đóng gói AAV. Thông thường, vectơ rAAV như được mô tả trong bản mô tả này thiếu trình tự gen đóng gói AAV bất kỳ để làm giảm tới mức tối thiểu khả năng mà AAV có khả năng sao chép được tạo ra bằng cách tái tổ hợp giữa gen đóng gói AAV và vectơ rAAV chỉ định. Theo nhiều phương án, các chế phẩm vectơ rAAV như được mô tả trong bản mô tả này là các chế phẩm chứa rất ít, nếu có, AAV có khả năng sao chép (rcAAV, còn được gọi là RCA) (ví dụ, ít hơn khoảng 1 rcAAV trên mỗi 10^2 hạt rAAV, ít hơn khoảng 1 rcAAV rên mỗi 10^4 hạt rAAV, ít hơn khoảng 1 rcAAV trên mỗi 10 hạt rAAV, ít hơn khoảng 1 rcAAV trên mỗi 10^{12} hạt rAAV, hoặc không có rcAAV).

Thuật ngữ "polynucleotit" dùng để chỉ dạng polyme của nucleotit có chiều dài

bất kỳ, bao gồm đeoxyribonucleotit hoặc ribonucleotit, hoặc chất tương tự của chúng. Polynucleotit có thể chứa nucleotit được biến đổi, như nucleotit được metyl hóa và chất tương tự nucleotit, và có thể bị gián đoạn bởi các thành phần không phải nucleotit. Nếu có, có thể biến đổi cấu trúc nucleotit trước hoặc sau khi lắp ráp polyme. Thuật ngữ polynucleotit, như được sử dụng trong bản mô tả này, dùng để chỉ phân tử sợi đơn và sợi đôi. Trừ khi có quy định khác cụ thể hoặc bắt buộc khác, bất kỳ phương án nào trong bản mô tả này bao gồm polynucleotit bao gồm cả dạng sợi đôi và mỗi trong số hai dạng sợi đơn bổ sung đã biết hoặc được dự đoán tạo ra dạng sợi đôi.

Polynucleotit hoặc polypeptit có tỷ lệ phần trăm "mức tương đồng trình tự" nhất định với polynucleotit hoặc polypeptit khác, nghĩa là, khi được sắp thẳng hàng, thì tỷ lệ phần trăm của bazơ hoặc axit amin giống nhau khi so sánh hai trình tự. Độ tương đồng trình tự có thể được xác định theo một số cách khác nhau. Để xác định mức tương đồng trình tự, các trình tự có thể được sắp thẳng hàng bằng cách sử dụng các phương pháp và chương trình máy tính, bao gồm BLAST, sẵn có trên toàn cầu tại trang web ncbi.nlm.nih.gov/BLAST/. Một thuật toán sắp thẳng hàng khác là FASTA, sẵn có trong Genetics Computing Group (GCG) package, từ Madison, Wisconsin, Mỹ, một công ty con của Oxford Molecular Group, Inc. Một phương pháp sắp thẳng hàng khác nữa được mô tả trong Methods in Enzymology, vol. 266: Computer Methods for Macromolecular Sequence Analysis (1996), ed. Doolittle, Academic Press, Inc., một bộ phận của Harcourt Brace & Co., San Diego, California, Mỹ. Một thuật toán được đặc biệt quan tâm là chương trình sắp thẳng hàng cho phép các khoảng trống trong trình tự. Smith- Waterman là một trong các thuật toán cho phép các khoảng trống trong sắp thẳng hàng trình tự. Xem Meth. Mol. Biol. 70: 173-187 (1997). Ngoài ra, chương trình GAP sử dụng phương pháp sắp thẳng hàng Needleman và Wunsch có thể được sử dụng để sắp thẳng hàng các trình tự. Xem J. Mol. Biol. 48: 443- 453 (1970).

Thuật ngữ "gen" dùng để chỉ polynucleotit thực hiện một số chức năng trong tế bào. Ví dụ, gen có thể chứa khung đọc mở có khả năng mã hóa sản phẩm gen. Một ví dụ về sản phẩm gen là protein, được phiên mã và dịch mã từ gen. Một ví dụ khác về sản phẩm gen là ARN, ví dụ, sản phẩm ARN chức năng, ví dụ, aptame, ARN can thiệp, ARN ribosom (rARN), ARN vận chuyển (tARN), ARN không mã hóa (ncARN), ARN dẫn hướng cho nucleaza, v..v., được phiên mã nhưng không dịch mã.

Thuật ngữ "sản phẩm biểu hiện gen" hoặc "sản phẩm gen" là phân tử thu được từ sự biểu hiện của gen cụ thể, như được xác định trên đây. Sản phẩm biểu hiện gen bao gồm, ví dụ, polypeptit, aptame, ARN can thiệp, ARN thông tin (mARN), rARN, tARN, ARN không mã hóa (ncRNA), và các sản phẩm tương tự.

Thuật ngữ "tác nhân siARN" ("ARN can thiệp nhỏ" hoặc "ARN can thiệp ngắn" (hoặc siARN)) là ARN sợi đôi chứa nucleotit hướng đích đến gen quan tâm ("gen đích"). "ARN sợi đôi" dùng để chỉ cấu trúc được tạo ra bằng cách bắt cặp bổ sung giữa hai vùng của phân tử ARN, tạo ra vùng ARN sợi đôi (dsARN). siARN được "hướng đích" đến gen mà trình tự nucleotit của phần siARN sợi đôi bổ sung với trình tự nucleotit của gen đích. Theo một số phương án, chiều dài của sợi đôi của siARN nhỏ hơn 30 nucleotit. Theo một số phương án, sợi đôi có thể có chiều dài 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11 hoặc 10 nucleotit. Theo một số phương án, chiều dài của sợi đôi là 19-25 nucleotit. Theo một số phương án, việc hướng đích gen qua trung gian siARN được thực hiện bằng việc sử dụng can thiệp ARN trực tiếp bởi ADN (ddARNi) là kỹ thuật làm im lặng gen sử dụng cấu trúc ADN để hoạt hóa con đường can thiệp ARN nội sinh của tế bào động vật. Cấu trúc ADN này được thiết kế để biểu hiện ARN sợi đôi tự bổ sung, thường là ARN kẹp tóc ngắn (shARN), mà khi được xử lý sẽ làm im lặng gen hoặc các gen đích. ARN bất kỳ, bao gồm mARN nội sinh hoặc ARN virut, có thể được làm im lặng bằng cách thiết kế các cấu trúc để biểu hiện ARN sợi đôi bổ sung với đích mARN mong muốn. Như vậy, phần ARN kép của tác nhân siARN có thể là một phần của cấu trúc kẹp tóc ngắn được gọi là shARN. Ngoài phần kép, cấu trúc kẹp tóc có thể chứa phần vòng nằm giữa hai trình tự tạo ra trình tự kép. Vòng có thể thay đổi về chiều dài. Theo một số phương án, vòng có chiều dài 5, 6, 7, 8, 9, 10, 11, 12 hoặc 13 nucleotit. Cấu trúc kẹp tóc cũng có thể chứa các phần nhô ra ở đầu 3' hoặc 5'. Theo một số phương án, phần nhô ra là phần nhô ra ở đầu 3' hoặc 5' có chiều dài 0, 1, 2, 3, 4 hoặc 5 nucleotit. Nói chung, hàm lượng của phẩm biểu hiện (ví dụ, mARN, polypeptit, v..v.) trong gen đích bị giảm đi bởi tác nhân siARN (ví dụ, siARN, shARN, v..v.) chứa trình tự nucleotit sợi đôi bổ sung đặc hiệu với ít nhất một đoạn có chiều dài 19-25 nucleotit (ví dụ, trình tự có chiều dài 20-21 nucleotit) của sản phẩm phiên mã gen đích, bao gồm vùng không được dịch mã ở đầu 5' (UT), ORF, hoặc vùng UT ở đầu 3'. Theo một số phương án, ARN

can thiệp ngắn có chiều dài khoảng 19-25 nucleotit. Xem, ví dụ, Công bố đơn quốc tế số WOO/44895, W099/32619, WO01/75164, WO01/92513, WO01/29058, WO01/89304, WO02/16620, và WO02/29858; và Công bố đơn yêu cầu cấp bằng sáng chế của Mỹ số 20040023390 mô tả về công nghệ siARN. SiARN và/hoặc shARN có thể được mã hóa bởi trình tự axit nucleic, và trình tự axit nucleic cũng có thể bao gồm vùng gen khởi đầu. Trình tự axit nucleic cũng có thể bao gồm tín hiệu polyađenyl hóa. Theo một số phương án, tín hiệu polyađenyl hóa là tín hiệu polyađenyl hóa tối thiểu tổng hợp.

Thuật ngữ "ARN đối nghĩa" bao gồm ARN bổ sung với sản phẩm biểu hiện gen. Ví dụ, ARN đối nghĩa hướng đích đến mARN đặc hiệu là tác nhân trên cơ sở ARN (hoặc có thể là ARN được biến đổi) cái mà bổ sung với mARN, trong đó sự lai hóa của ARN đối nghĩa với mARN làm thay đổi sự biểu hiện của mARN (ví dụ, nhờ sửa đổi độ ổn định của ARN, sửa đổi sự dịch mã của ARN, v..v.). Cũng bao gồm trong "ARN đối nghĩa" là axit nucleic mã hóa ARN đối nghĩa.

Liên quan đến "tác nhân CRISPR/Cas9", thuật ngữ "CRISPR" bao gồm trình tự lặp lại đoạn gập đôi ngắn có khoảng trống đều đặn được tạo cụm/hệ liên kết CRISPR (Cas) được phát triển để tạo ra vi khuẩn và cổ khuẩn có khả năng miễn dịch thích nghi chống lại virut và plasmit bằng cách sử dụng ARN CRISPR (crARN) để dẫn hướng quá trình làm im lặng axit nucleic đang lan tràn. Protein Cas9 (hoặc dạng tương đương về chức năng và/hoặc biến thể của chúng, tức là, protein giống Cas9) chứa hoạt tính ADN enđonucleaza tự nhiên tùy thuộc vào sự liên kết của protein với hai phân tử ARN tự nhiên hoặc tổng hợp được gọi là crARN và tracrARN (còn được gọi là ARN dẫn hướng). Trong một số trường hợp, hai phân tử được liên kết cộng hóa trị để tạo ra một phân tử duy nhất (còn được gọi là ARN dẫn hướng đơn ("sgARN")). Do đó, Cas9 hoặc protein giống Cas9 được liên kết với ARN hướng đích ADN (thuật ngữ này bao gồm cả cấu trúc ARN dẫn hướng hai phân tử và cấu trúc ARN dẫn hướng một phân tử), hoạt hóa Cas9 hoặc protein giống Cas9 và dẫn hướng protein đến trình tự axit nucleic đích.

Nếu Cas9 hoặc protein giống Cas9 giữ lại chức năng enzym tự nhiên của nó, nó sẽ phân cắt ADN đích để tạo ra đứt gãy sợi đôi, có thể dẫn đến việc sửa đổi hệ gen (tức là, chỉnh sửa: mất, thêm (khi có mặt polynucleotit cho), thay thế, v..v.), nhờ đó thay

đổi sự biểu hiện của gen. Một số biến thể của Cas9 (các biến thể này cũng nằm trong thuật ngữ giống Cas9) đã được sửa đổi sao cho chúng có hoạt tính phân cắt ADN giảm đi (trong một số trường hợp, chúng phân cắt sợi đơn thay vì cả hai sợi ADN đích, trong khi trong các trường hợp khác, chúng không có hoặc có hoạt tính phân cắt ADN giảm mạnh). Protein giống Cas9 có hoạt tính phân cắt ADN giảm (thậm chí không có hoạt tính phân cắt ADN) có thể vẫn được dẫn hướng đến ADN đích để phong bế hoạt tính ARN polymeaza. Theo cách khác, Cas9 hoặc protein giống Cas9 có thể được biến đổi bằng cách dung hợp miền hoạt hóa phiên mã VP64 với protein Cas9 và vận chuyển đồng thời protein dung hợp với protein trợ giúp MS2-P65-HSF1 và ARN dẫn hướng đơn chứa aptame MS2 ARN ở vòng bốn và vòng gốc để tạo ra phức hợp trung gian hoạt hóa tổng hợp (Cas9-SAM) trong tế bào hoạt hóa quá trình phiên mã. Do đó, protein giống Cas9 được làm bất hoạt bằng enzym này có thể hướng đích đến vị trí nhất định trong ADN đích nhờ ARN hướng đích ADN để phong bế hoặc hoạt hóa quá trình phiên mã của ADN đích. Thuật ngữ "tác nhân CRISPR/Cas9" như được sử dụng trong bản mô tả này bao gồm tất cả các dạng CRISPR/Cas9 như đã mô tả ở trên hoặc đã được biết đến trong lĩnh vực kỹ thuật.

Thông tin chi tiết liên quan đến tác nhân CRISPR có thể được tìm thấy, ví dụ, trong Jinek et. al., Science. 2012 Aug 17;337(6096):816-21: "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity"; (b) Qi et al., Cell. 2013 Feb 28; 152(5): 1173-83: "Repurposing CRISPR as an RNA- guided platform for sequence- specific control of gene expression", và (c) Đơn yêu cầu cấp bằng sáng chế của Mỹ số 13/842,859 và Đơn quốc tế số PCT/US13/32589; nội dung của các tài liệu này được đưa vào bản mô tả bằng cách viện dẫn. Do đó, thuật ngữ "tác nhân CRISPR" như được sử dụng trong bản mô tả này bao gồm tác nhân bất kỳ (hoặc axit nucleic mã hóa tác nhân như vậy), bao gồm các trình tự tự nhiên và/hoặc tổng hợp, mà có thể được sử dụng trong hệ thống trên cơ sở Cas9 (ví dụ, Cas9 hoặc protein giống Cas9; thành phần bất kỳ của ARN hướng đích ADN, ví dụ, ARN giống crARN, ARN giống tracrARN, ARN dẫn hướng đơn, v.v.; polynucleotit cho; và các tác nhân tương tự).

Thuật ngữ "nucleaza ngón tay kẽm" (ZFN) nghĩa là ADN endonucleaza nhân tạo được tạo ra bằng cách dung hợp miền liên kết ADN ngón tay kẽm với miền phân cắt ADN. ZFN có thể được thiết kế để hướng đích các trình tự ADN mong muốn và

điều này có thể làm cho nucleaza ngón tay kẽm phân cắt các trình tự đích duy nhất. Khi được đưa vào tế bào, ZFN có thể được sử dụng để chỉnh sửa ADN đích trong tế bào (ví dụ, hệ gen của tế bào) bằng cách tạo ra đứt gãy sợi đôi. Để biết thêm thông tin về việc sử dụng ZFN, xem, ví dụ: Asuri et al., Mol Ther. 2012 Feb;20(2):329-38; Bibikova et al. Science. 2003 May 2;300(5620):764; Wood et al. Science. 2011 Jul 15;333(6040):307; Ochiai et al. Genes Cells. 2010 Aug;15(8):875-85; Takasu et. al., Insect Biochem Mol Biol. 2010 Oct;40(10):759-65; Ekker et al, Zebrafish 2008 Summer;5(2): 121-3; Young et al, Proc Natl Acad Sci U S A. 2011 Apr 26;108(17):7052-7; Goldberg et al, Cell. 2010 Mar 5;140(5):678-91; Geurts et al, Science. 2009 Jul 24;325(5939):433; Flisikowska et al, PLoS One. 2011;6(6):e21045. doi: 10.1371/journal.pone.0021045. Epub 2011 Jun 13; Hauschild et al, Proc Natl Acad Sci U S A. 2011 Jul 19;108(29): 12013-7; và Yu et al, Cell Res. 2011 Nov;21(l 1): 1638-40; nội dung của các tài liệu này được đưa vào bản mô tả bằng cách viện dẫn về nội dung liên quan đến ZFN. Thuật ngữ "tác nhân ZFN" bao gồm nucleaza ngón tay kẽm và/hoặc polynucleotit bao gồm trình tự nucleotit mã hóa nucleaza ngón tay kẽm.

Thuật ngữ “Nucleaza tác động giống chất hoạt hóa quá trình phiên mã” hoặc tác nhân “TALEN” dùng để chỉ nucleaza tác động giống chất hoạt hóa quá trình phiên mã (TALEN). TALEN là enđonucleaza ADN nhân tạo được tạo ra bằng cách dung hợp miền liên kết ADN tác động TAL (giống chất hoạt hóa quá trình phiên mã) với miền đứt gãy ADN. TALEN có thể nhanh chóng thiết kế để liên kết thực tế trình tự ADN mong muốn bất kỳ và khi được đưa vào tế bào, TALEN có thể được sử dụng để chỉnh sửa ADN đích trong tế bào (ví dụ, hệ gen của tế bào) bằng cách tạo ra đứt gãy sợi đôi. Để biết thêm thông tin về việc sử dụng TALEN, xem, ví dụ: Hockemeyer et al. Nat Biotechnol. 2011 Jul 7;29(8):731-4; Wood et al. Science. 2011 Jul 15;333(6040):307; Tesson et al. Nat Biotechnol. 2011 Aug 5;29(8):695-6; và Huang et. al., Nat Biotechnol. 2011 Aug 5;29(8):699-700; nội dung của các tài liệu này được đưa vào bản mô tả bằng cách viện dẫn về nội dung liên quan đến TALEN. Thuật ngữ "tác nhân TALEN " bao gồm TALEN và/hoặc polynucleotit bao gồm trình tự nucleotit mã hóa TALEN.

Thuật ngữ "yếu tố kiểm soát" hoặc "trình tự kiểm soát" dùng để chỉ trình tự nucleotit liên quan đến sự tương tác của phân tử góp phần vào việc điều hòa chức năng

của polynucleotit, bao gồm sao chép, nhân đôi, phiên mã, cắt ghép, dịch mã, hoặc biến tính polynucleotit. Quá trình điều hòa có thể tác động đến tần suất, tốc độ, hoặc độ đặc hiệu của quá trình, và có thể tăng cường hoặc ức chế về bản chất. Yếu tố kiểm soát đã biết trong lĩnh vực kỹ thuật này bao gồm, ví dụ, các trình tự điều hòa phiên mã như vùng gen khởi đầu và vùng tăng cường. Vùng gen khởi đầu là vùng ADN trong các điều kiện nhất định có khả năng liên kết ARN polymeaza và khơi mào phiên mã vùng mã hóa thường đặt xuôi chiều (theo hướng đầu 3') từ vùng gen khởi đầu. Vùng gen khởi đầu có thể tác động phổ biến, tức là hoạt động ở nhiều loại tế bào, ví dụ, vùng gen khởi đầu CAG hoặc CMV; hoặc mô hoặc tế bào đặc hiệu, ví dụ, vùng gen khởi đầu đặc hiệu mô trong tế bào cơ tim.

Thuật ngữ "được liên kết một cách có điều khiển" hoặc "được liên kết có điều khiển" dùng để chỉ vị trí kề nhau của yếu tố di truyền, trong đó các yếu tố này ở trong mối quan hệ cho phép chúng hoạt động theo cách mong muốn. Ví dụ, vùng gen khởi đầu được liên kết một cách có điều khiển với vùng mã hóa nếu vùng gen khởi đầu trợ giúp khơi mào quá trình phiên mã trình tự mã hóa. Có thể có các gốc can thiệp giữa vùng gen khởi đầu và vùng mã hóa chừng nào mối quan hệ chức năng này được duy trì.

Thuật ngữ "vectơ biểu hiện" bao gồm vectơ bao gồm vùng polynucleotit mã hóa polypeptit quan tâm, và được sử dụng để tác động đến sự biểu hiện của protein trong tế bào đích dự tính. Vectơ biểu hiện cũng bao gồm yếu tố kiểm soát được liên kết một cách có điều khiển với vùng mã hóa để thuận lợi biểu hiện protein ở đích. Việc kết hợp các yếu tố kiểm soát và một gen hoặc nhiều gen mà liên kết có điều khiển với nhau để biểu hiện đôi khi được gọi là "băng biểu hiện", một lượng lớn chúng là đã biết và sẵn có trong lĩnh vực kỹ thuật này hoặc có thể được tạo cấu trúc một cách dễ dàng từ các thành phần sẵn có trong lĩnh vực kỹ thuật này

Thuật ngữ "khác loại" nghĩa là có nguồn gốc từ thực thể khác biệt về kiểu gen so với phần còn lại của thực thể mà nó được so sánh. Ví dụ, polynucleotit được đưa bằng kỹ thuật thiết kế di truyền vào trong plasmit hoặc vectơ có nguồn gốc từ loài khác là polynucleotit khác loại. Vùng gen khởi đầu được loại ra khỏi trình tự mã hóa nguyên bản của nó và được liên kết một cách có điều khiển với trình tự mã hóa mà không tìm thấy trong tự nhiên có liên kết với là vùng gen khởi đầu khác loại. Do đó, ví

dụ, rAAV bao gồm trình tự axit nucleic khác loại mã hóa cho sản phẩm gen khác loại là rAAV bao gồm polynucleotit bình thường không nằm trong AAV có trong tự nhiên, kiểu dại, và sản phẩm gen khác loại đã được mã hóa là sản phẩm gen bình thường không được mã hóa bởi AAV kiểu dại có trong tự nhiên.

Thuật ngữ "biến đổi di truyền" và "biến đổi di truyền" (và các biến thể ngữ pháp của chúng), được sử dụng hoán đổi cho nhau trong bản mô tả này để chỉ quy trình trong đó yếu tố di truyền (ví dụ, polynucleotit) được đưa vào tế bào ngoài con đường nguyên phân và giảm phân. Yếu tố này có thể khác loại với tế bào, hoặc nó có thể là bản sao bổ sung hoặc phiên bản cải thiện của yếu tố đã có mặt trong tế bào. Việc biến đổi di truyền có thể được thực hiện, ví dụ, bằng cách tải nạp tế bào với plasmit tái tổ hợp hoặc polynucleotit khác bằng quy trình bất kỳ đã biết trong lĩnh vực kỹ thuật, như điện di, kết tủa canxi phosphat, hoặc tiếp xúc với phức hợp polynucleotit-liposom. Biến đổi di truyền cũng có thể được thực hiện, ví dụ, bằng cách biến nạp hoặc lây nhiễm bằng ADN hoặc ARN virut hoặc vectơ virut. Nói chung, yếu tố di truyền được đưa vào nhiễm sắc thể hoặc tiểu nhiễm sắc thể trong tế bào; nhưng biến đổi bất kỳ làm thay đổi kiểu hình và/hoặc kiểu gen của tế bào và thế hệ con của nó cũng nằm trong thuật ngữ này.

Liên quan đến việc biến đổi tế bào, thuật ngữ "được biến đổi di truyền" hoặc "được biến nạp" hoặc "được tải nạp" hoặc “được tải nạp” bởi ADN ngoại sinh (ví dụ, nhờ virut tái tổ hợp) dùng để chỉ khi ADN này đã được đưa vào bên trong tế bào. Sự có mặt của ADN ngoại sinh dẫn đến việc thay đổi di truyền tạm thời hoặc vĩnh viễn. ADN biến nạp có thể hoặc không thể được tích hợp (được liên kết cộng hóa trị) vào trong hệ gen của tế bào. "Dòng vô tính" là quần thể tế bào có nguồn gốc từ chỉ một tế bào hoặc tổ tiên chung nhờ nguyên phân. "Dòng tế bào" là dòng tế bào vô tính của tế bào chưa biệt hóa có khả năng sinh trưởng ổn định *in vitro* cho nhiều thế hệ.

Như được sử dụng trong bản mô tả này, tế bào được cho là được sửa đổi, được tải nạp, được biến đổi di truyền, hoặc được biến nạp “ổn định” với trình tự gen nếu trình tự này là sẵn có để thực hiện chức năng của nó trong quá trình nuôi cấy mở rộng tế bào *in vitro* và/hoặc cho mỗi giai đoạn thời gian mở rộng *in vivo*. Nói chung, tế bào này được biến đổi “di truyền” (được biến đổi di truyền) trong đó biến đổi gen được đưa vào cũng không thể di truyền được bởi thế hệ con của tế bào sửa đổi.

các thuật ngữ "polypeptit", "peptit", và "protein" được sử dụng thay thế cho nhau trong bản mô tả này để chỉ polyme của axit amin có chiều dài bất kỳ. Các thuật ngữ này cũng bao gồm polyme axit amin đã được biến đổi; ví dụ, hình thành liên kết đisulfua, glycosyl hóa, lipit hóa, phosphoryl hoá, hoặc liên hợp với thành phần đánh dấu. Polypeptit như polypeptit chống hình thành mạch, polypeptit bảo vệ thần kinh, và các polypeptit tương tự, khi được đề cập trong phần nội dung về vận chuyển sản phẩm gen đến đối tượng động vật có vú, và các chế phẩm của chúng, dùng để chỉ polypeptit bẩm sinh tương ứng, hoặc mảnh bất kỳ hoặc dẫn xuất được thiết kế di truyền của chúng, mà vẫn giữ được chức năng sinh hóa mong muốn của protein bẩm sinh. Tương tự, liên quan đến axit nucleic mã hóa polypeptit chống hình thành mạch, axit nucleic mã hóa polypeptit bảo vệ thần kinh, và axit nucleic khác như vậy dùng để vận chuyển sản phẩm gen đến đối tượng động vật có vú (mà có thể được gọi là "gen chuyển" để được vận chuyển đến tế bào nhận), bao gồm polynucleotit mã hóa polypeptit bẩm sinh hoặc mảnh bất kỳ hoặc dẫn xuất được thao tác di truyền có chức năng sinh hóa mong muốn.

Như được sử dụng trong bản mô tả này, plasmit, axit nucleic, vectơ, virut, virion, tế bào chủ, protein, hoặc cơ chất khác "được phân lập" dùng để chỉ chế phẩm không có ít nhất một vài trong số thành phần còn lại mà cũng có thể có mặt tại nơi mà chất hoặc chất tương tự bát nguồn trong tự nhiên hoặc được điều chế ban đầu. Do đó, ví dụ, chất phân lập có thể được tạo ra bằng cách sử dụng kỹ thuật tinh chế để làm giàu nó từ hỗn hợp nguồn. Việc làm giàu có thể được đo trên cơ sở tuyệt đối, như khối lượng cho mỗi thể tích của dung dịch, hoặc nó có thể được đo tính theo chất can thiệp tiềm năng thứ hai có mặt trong hỗn hợp nguồn. Việc làm tăng việc làm giàu theo các phương án theo sáng chế đang được phân lập ngày càng gia tăng. Plasmit, axit nucleic, vectơ, virut, tế bào chủ, hoặc chất khác phân lập theo một số phương án được tinh chế, ví dụ, từ khoảng 80% đến khoảng 90% tinh khiết, ít nhất khoảng 90% tinh khiết, ít nhất khoảng 95% tinh khiết, ít nhất khoảng 98% tinh khiết, hoặc ít nhất khoảng 99%, hoặc nhiều hơn.

Trong bản mô tả này, các thuật ngữ "việc điều trị", "điều trị", và các thuật ngữ tương tự, dùng để chỉ việc thu được hiệu quả dược lý và/hoặc sinh lý học mong muốn. Hiệu quả này có thể là phòng bệnh liên quan đến việc ngăn ngừa hoàn toàn hoặc một

phần bệnh lý hoặc triệu chứng của chúng và/hoặc có thể là điều trị liên quan đến việc chữa trị một phần hoặc hoàn toàn bệnh và/hoặc tác động bất lợi do bệnh gây ra. "Điều trị", như được sử dụng trong bản mô tả này, bao gồm việc điều trị bệnh lý bất kỳ ở động vật có vú, cụ thể là ở người, và bao gồm: (a) ngăn ngừa bệnh (và/hoặc các triệu chứng do bệnh lý gây ra) xảy ra ở đối tượng mà mắc bệnh hoặc nguy cơ mắc bệnh nhưng chưa được chẩn đoán khi có nó; (b) ức chế bệnh (và/hoặc các triệu chứng do bệnh gây ra), tức là, làm ngừng sự phát triển của nó; và (c) làm thuyên giảm bệnh (và/hoặc các triệu chứng do bệnh gây ra), tức là, gây ra sự thoái biến của bệnh (và/hoặc các triệu chứng do bệnh lý gây ra), tức là, cải thiện bệnh và/hoặc một hoặc nhiều triệu chứng của bệnh. Ví dụ, chế phẩm và phương pháp theo sáng chế có thể hướng đến việc điều trị bệnh cơ. Các phương pháp không giới hạn để đánh giá bệnh cơ và điều trị nó bao gồm đo việc sản xuất protein trị liệu (ví dụ, sinh thiết cơ sau đó hóa mô miễn dịch hoặc lấy mẫu huyết thanh sau đó ELISA hoặc thử nghiệm hoạt tính enzym), đo triệu chứng của suy tim (ví dụ, phân loại chức năng của Hiệp hội Tim mạch New York hoặc Câu hỏi về Cuộc sống với Bệnh suy tim ở tiểu bang Minnesota), tình trạng chức năng tim (ví dụ, thử nghiệm đi bộ 6 phút hoặc tiêu thụ oxi tối đa), phân tích chất đánh dấu sinh học (ví dụ peptide natriuretic não prohormone đầu N), chức năng / tổ chức lại tâm thất trái (ví dụ phân suất tống máu tâm thất trái hoặc thể tích tâm thất trái thu), sức mạnh cơ bắp (ví dụ như Hội đồng nghiên cứu y tế Điều tra lâm sàng bệnh loạn dưỡng cơ Duchenne, động lực học cầm tay hoặc nâng tạ tối đa), chức năng cơ bắp (ví dụ Thang đo Vignos, Kiểm tra chức năng theo thời gian, Điểm số khả năng vận động của Hammersmith, thời gian đứng lên từ sàn, thử nghiệm đi bộ, Thang đo chức năng vận động, Đánh giá sự đi lại North Star, Thử nghiệm 9 lỗ, hoặc Thử nghiệm của Bệnh viện nhi Philadelphia về rối loạn thần kinh cơ), các triệu chứng bệnh cơ (ví dụ: thang điểm triệu chứng thần kinh cơ hoặc đáp ứng chung lâm sàng), chức năng ty thể (ví dụ từ quang phổ cộng hưởng ^{31}P), đánh giá dựa trên bảng câu hỏi về chất lượng cuộc sống, kết quả báo cáo của bệnh nhân hoặc các hoạt động hàng ngày.

Các thuật ngữ "cá thể", "vật chủ", "đối tượng", và "bệnh nhân" được sử dụng thay thế cho nhau trong bản mô tả này, và dùng để chỉ động vật có vú, bao gồm, nhưng không chỉ giới hạn ở, người; linh trưởng không phải người, bao gồm khỉ; động vật có vú dùng trong thể thao (ví dụ, ngựa); động vật có vú nuôi trong trang trại (ví dụ, cừu,

dê, v..v.); động vật nuôi có vú nuôi làm cảnh (chó, mèo, v..v.); và loài gặm nhấm (ví dụ, chuột nhắt, chuột cống, v..v.).

Theo một số phương án, cá thể là người trước đây đã từng phơi nhiễm tự nhiên với AAV và dẫn đến việc chứa kháng thể kháng AAV (tức là, kháng thể trung hòa AAV). Theo một số phương án, cá thể là người đã từng được dùng vectơ AAV trước đó (và dẫn đến có thể chứa kháng thể kháng AAV) và cần dùng lại vectơ để điều trị tình trạng bệnh lý khác hoặc để điều trị tiếp các tình trạng bệnh lý này. Trên cơ sở các kết quả tích cực trong các thử nghiệm lâm sàng liên quan đến việc vận chuyển gen AAV đến, ví dụ, gan, cơ, và võng mạc - tất cả các mô bị tác động bởi kháng thể trung hòa chống lại chất dẫn thuốc này – có rất nhiều ứng dụng điều trị/đích bệnh.

Thuật ngữ "lượng có hiệu quả" như được sử dụng trong bản mô tả này là lượng đủ để tạo ra kết quả lâm sàng có lợi hoặc mong muốn. Lượng có hiệu quả có thể được dùng trong một hoặc nhiều lần dùng. Đối với mục đích của bản mô tả này, lượng có hiệu quả của hợp chất (ví dụ, virion rAAV lây nhiễm) là lượng đủ để làm giảm, cải thiện, ổn định, đảo ngược, ngăn chặn, làm chậm hoặc trì hoãn tiến triển của (và/hoặc các triệu chứng liên quan đến) tình trạng bệnh lý cụ thể (ví dụ, bệnh võng mạc). Do đó, lượng có hiệu quả của virion rAAV lây nhiễm là lượng virion rAAV lây nhiễm mà có thể vận chuyển hiệu quả axit nucleic khác loại đến tế bào đích (hoặc các tế bào đích) của cá thể. Lượng có hiệu quả có thể được xác định tiền lâm sàng bằng cách, ví dụ, dò trong tế bào hoặc mô, sản phẩm gen (ARN, protein) được mã hóa bởi trình tự axit nucleic khác loại bằng cách sử dụng kỹ thuật đã biết trong lĩnh vực kỹ thuật này, ví dụ RT-PCR, thấm tách miễn dịch, ELISA, huỳnh quang hoặc các thiết bị hiện dữ liệu báo cáo khác, và các kỹ thuật tương tự. Lượng có hiệu quả có thể được xác định lâm sàng bằng, ví dụ, phát hiện sự thay đổi trong giai đoạn khởi phát hoặc tiến triển của bệnh bằng cách sử dụng các phương pháp đã biết trong lĩnh vực kỹ thuật này, ví dụ, thử nghiệm đi bộ 6 phút, phân suất tống máu tâm thất trái, đo lực cầm tay, Thang đo Vignos, và các kỹ thuật tương tự, như được đề cập trong bản mô tả này và đã được biết đến trong lĩnh vực kỹ thuật này.

Thuật ngữ “tế bào cơ” hoặc “mô cơ” trong bản mô tả này dùng để chỉ tế bào hoặc nhóm tế bào có nguồn gốc từ cơ của cơ quan bất kỳ, bao gồm, nhưng không chỉ giới hạn ở, tế bào cơ xương, cơ tim, cơ trơn (ví dụ, từ đường tiêu hóa, bọng đái và

mạch máu) và cơ hoành. Các tế bào cơ này có thể được biệt hóa hoặc không được biệt hóa chẳng hạn như nguyên bào cơ, tế bào cơ, ống cơ, tế bào cơ tim, và nguyên bào cơ tim. Do mô cơ dễ dàng tiếp cận với hệ tuần hoàn, protein được sản xuất và tiết ra từ các tế bào cơ và mô *in vivo* sẽ đi vào máu một cách hợp lý để mang lại tác dụng toàn thân, từ đó cung cấp mức độ trị liệu, duy trì liên tục của protein tiết ra từ cơ.

Thuật ngữ "tiến hóa trực hệ" dùng để chỉ phương pháp thao tác di truyền capsit, *in vitro* và/hoặc *in vivo*, cạnh tranh với tiến hóa tự nhiên qua các vòng đa dạng hóa di truyền lặp lại và các quy trình chọn lọc, nhờ đó tích lũy đột biến có lợi cải thiện dần dần chức năng của phân tử sinh học. Tiến hóa trực hệ mang lại sự tiến hóa của phương pháp *in vivo* được gọi là "biopanning" để chọn lọc biến thể AAV từ thư viện mà các biến thể có mức lây nhiễm hiệu quả hơn đối với tế bào hoặc dạng mô quan tâm.

Virut adeno kết hợp (AAV) là họ virut parvo với hệ gen ADN sợi đơn 4,7 kb nằm bên trong capsit không vỏ. Hệ gen virut của AAV tự nhiên có 2 trình tự đầu lặp lại đảo ngược (inverted terminal repeat - ITR) – hoạt động chức năng như điểm khởi đầu sao chép và tín hiệu đóng gói của virut – nằm bên sườn 2 khung đọc mở (ORF) nguyên thủy : *rep* (mã hóa protein có chức năng trong sao chép virut, điều hòa phiên mã, tích hợp đặc hiệu vị trí, và lắp ráp virion) và *cap*. *Cap* ORF mã hóa cho 3 protein cấu trúc lắp ráp để tạo ra capsit virut 60-mer. Nhiều biến thể AAV và kiểu huyết thanh có trong tự nhiên đã được phân lập, và không có biến thể AAV và kiểu huyết thanh có liên quan đến bệnh của người.

Các phiên bản tái tổ hợp của AAV có thể được sử dụng làm vectơ vận chuyển gen, trong đó gen đánh dấu hoặc gen trị liệu quan tâm được cài vào giữa các ITR ở vị trí của *rep* và *cap*. Các vectơ này được tạo ra để tải nạp cả các tế bào đang phân chia và không phân chia *in vitro* và *in vivo* và có thể dẫn đến biểu hiện gen chuyển một cách ổn định trong nhiều năm trong mô hậu nguyên phân. Xem, ví dụ, Knipe DM, Howley PM. *Fields' Virology*. Lippincott Williams & Wilkins, Philadelphia, PA, Mỹ, 2007; Gao G-P, Alvira MR, Wang L, Calcedo R, Johnston J, Wilson JM. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc Natl Acad Sci Mỹ* 2002; 99: 11854–9;Atchison RW, Casto BC, Hammon WM. Adenovirus-Associated Defective Virus Particles. *Science* 1965; 149: 754–6; Hoggan MD, Blacklow NR, Rowe WP. Studies of small DNA viruses found in various adenovirus

preparations: physical, biological, và immunological characteristics. *Proc Natl Acad Sci Mỹ* 1966; 55: 1467–74; Blacklow NR, Hoggan MD, Rowe WP. Isolation of adenovirus-associated viruses from man. *Proc Natl Acad Sci Mỹ* 1967; 58: 1410–5; Bantel-Schaal U, zur Hausen H. Characterization of the DNA of a defective human parvovirus isolated from a genital site. *Virology* 1984; 134: 52–63; Mayor HD, Melnick JL. Small deoxyribonucleic axit-containing viruses (picodnavirus group). *Nature* 1966; 210: 331–2; Mori S, Wang L, Takeuchi T, Kanda T. Two novel adeno-associated viruses from cynomolgus monkey: pseudotyping characterization of capsit protein. *Virology* 2004; 330: 375–83; Flotte TR. Gene therapy progress và prospects: recombinant adeno-associated virus (rAAV) vectors. *Gene Ther* 2004; 11: 805–10.

AAV tái tổ hợp (trong bản mô tả này đơn giản được gọi là "AAV") đã mang lại kết quả đầy hứa hẹn trong số lượng thử nghiệm lâm sàng ngày càng tăng. Tuy nhiên, có những trở ngại đối với việc vận chuyển gen có thể hạn chế tính hữu dụng của AAV, như đáp ứng miễn dịch chống lại capsit, tải nạp thấp trong các mô nhất định, không có khả năng vận chuyển đích đến các loại tế bào cụ thể và khả năng mang tương đối thấp. Trong nhiều trường hợp, không có đủ kiến thức về cơ học để thiết kế hiệu quả hợp lý với khả năng cải thiện AAV. Thay vào đó, sự tiến hóa trực hệ đã xuất hiện như một chiến lược để tạo ra các biến thể AAV mới đáp ứng các nhu cầu y sinh cụ thể. Tiến hóa trực hệ đưa ra chiến lược khai thác sự đa dạng hóa và các quy trình chọn lọc di truyền để có thể tích tụ được đột biến có lợi cải thiện dần dần chức năng của phân tử sinh học. Trong quy trình này, gen AAV *cap* kiểu dại được đa dạng hóa bằng một vài phương pháp để tạo ra thư viện di truyền lớn được đóng gói để tạo ra thư viện của hạt virut, và sau đó đặt áp lực chọn lọc để phân lập các biến thể mới có thể vượt qua được rào cản vận chuyển gen. Quan trọng là, cơ sở cơ học củavấn đề vận chuyển gen không cần thiết phải đượcbiết đến với sự tiến hóa trực hệ của chức năng, do đó có thể thúc đẩy sự phát triển của các vectơ được tăng cường.

Thông thường, các biến thể được bộc lộ trong bản mô tả này được tạo ra bằng cách sử dụng thư viện và/hoặc các thư viện AAV. Thư viện hoặc các thư viện AAV này được tạo ra bằng cách gây đột biến gen *cap*, là gen mã hóa protein cấu trúc của capsit AAV, bằng nhiều kỹ thuật tiến hóa trực hệ đã biết và có sẵn đối với người có hiểu biết trong lĩnh vực thiết kế hệ gen virut. Xem, ví dụ, Bartel et al. Am. Soc. Gene

Cell Ther. 15th Annu. Meet. 20, S140 (2012); Bowles, D. et al. J. Virol. 77, 423-432 (2003); Gray et al. Mol. Ther. 18, 570-578 (2010); Grimm, D. et al. J. Virol. 82, 5887-5911; Koerber, J. T. et al. Mol. Ther. 16, 1703-1709 (2008); Li W. et al. Mol. Ther. 16, 1252-1260 (2008); Koerber, J. T. et al. Methods Mol. Biol. 434, 161-170 (2008); Koerber, J. T. et al. Hum. Gene Ther. 18, 367-378 (2007); và Koerber, J. T. et al. Mol. Ther. 17, 2088-2095 (2009). Các kỹ thuật này là, nhưng không chỉ giới hạn ở: i) PCR úp sấp sai hỏng (Error-prone PCR) để đưa đột biến điểm ngẫu nhiên vào khung đọc mở (ORF) *cap* AAV ở tỷ lệcó thể biến đổi, định trước; ii) tái tổ hợp virut hoặc "xáo trộn ADN" *in vitro* hoặc *in vivo* để tạo ra thể khảm ngẫu nhiên của gen *cap* AAV để thu được thư viện gen với nhiều kiểu huyết thanh AAV; iii) chèn peptit ngẫu nhiên tại vị trí capsit đã định bằng cách nối oligonucleotit suy biến trong ORF *cap*; iv) chèn các trình tự mã hóa peptit đã định vào các vị trí ngẫu nhiên của ORF *cap* AAV sử dụng gây đột biến gen nhảy; v) thay thế vòng bề mặt của capsit AAV bằng thư viện của các trình tự peptit được thiết kế về mặt thông tin sinh học dựa trên mức độ bảo toàn của mỗi vị trí axit amin trong số kiểu huyết thanh AAV tự nhiên và các biến thể để tạo ra các thư viện "trao đổi vòng"; vi) thay thế axit amin ngẫu nhiên ở các vị trí suy biến giữa các kiểu huyết thanh AAV để tạo ra thư viện của các biến thể tổ tiên (Santiago-Ortiz et al., 2015); và kết hợp các kỹ thuật này.

Xáo trộn ADN tạo ra thể khảm kết hợp được các đặc tính mẹ của chúng theo cách độc đáo và, thường có lợi; tuy nhiên, một số xáo trộn này có thể không có khả năng đóng gói mà, trong thực tế, làm giảm tính đa dạng của thư viện. Việc tập trung tính đa dạng của thư viện vào trong (các) vùng cụ thể của capsit đạt được nhờ kỹ thuật chèn peptit như, nhưng không chỉ giới hạn ở, các kỹ thuật iii-iv) nêu trên. Tính đa dạng của thư viện cũng được tập trung trong kỹ thuật như kỹ thuật nêu trong mục v) nêu trên, và sự tập trung này được hướng vào trong (các) vùng cụ thể của capsit bằng kỹ thuật chẳng hạn như kỹ thuật v) ở trên, và việc tập trung này hướng đến nhiều vùng siêu biến, nằm trên bề mặt vòng tiếp xúc, của capsit AAV. Trong khi nhiều kỹ thuật tạo ra capsit biến thể chỉ có vùng nhỏ capsit bị đột biến, các kỹ thuật này có thể bắt cặp với chiến lược gây đột biến bổ sung để biến đổi toàn bộ capsit.

Một khi thư viện hoặc các thư viện AAV được tạo ra, virut sau đó được đóng gói, sao cho mỗi hạt AAV nằm trong capsit đột biến bao quanh gen *cap* mã hóa cho

capsit này, và được tinh sạch. Các biến thể của thư viện sau đó được đưa vào kỹ thuật tạo áp lực chọn lọc *in vitro* và/hoặc *in vivo* đã biết và có sẵn đối với người có hiểu biết trong lĩnh vực AAV. Xem, ví dụ, Maheshri, N. et al. Nature Biotech. 24, 198-204 (2006); Dalkara, D. et al. Sci. Transl. Med. 5, 189ra76 (2013); Lisowski, L. et al. Nature. 506, 382-286 (2013); Yang, L. et al. PNAS. 106, 3946-3951 (2009); Gao, G. et al. Mol. Ther. 13, 77-87 (2006); và Bell, P. et al. Hum. Gene. Ther. 22, 985-997 (2011). Ví dụ, nhưng không chỉ giới hạn ở, biến thể AAV có thể được chọn lọc bằng cách sử dụng i) cột ái lực trong đó giải hấp các phân đoạn khác nhau thu được các biến thể có các đặc tính liên kết được biến đổi; ii) tế bào nguyên thủy – được phân lập từ các mẫu mô hoặc dòng tế bào bất tử bắt chước hành vi của tế bào trong cơ thể người – thu được biến thể AAV có hiệu lực và/hoặc tính đặc hiệu mô tăng lên; iii) mô hình động vật - bắt chước môi trường trị liệu gen lâm sàng - thu được biến thể AAV đã lây nhiễm thành công mô đích; iv) mô hình ghép ngoại lai người thu được biến thể AAV đã lây nhiễm tế bào người đã ghép; và/hoặc kết hợp của các kỹ thuật chọn lọc này.

Một khi virut đã được chọn, chúng có thể được thu hồi bằng kỹ thuật đã biết, nhưng không chỉ giới hạn ở, kỹ thuật sao chép qua trung gian virut ađeno, khuếch đại PCR, giải trình tự thế hệ tiếp theo và tạo dòng vô tính, và các kỹ thuật tương tự. Dòng vô tính virut sau đó được làm giàu qua các vòng chọn lọc lặp đi lặp lại và ADN AAV được phân lập để thu hồi gen *cap* biến thể quan tâm được chọn. Các biến thể được chọn này có thể được biến đổi hoặc gây đột biến tiếp và như vậy được coi như một sự khởi đầu mới cho các bước chọn lọc tiếp theo để làm tăng sự phù hợp của virut AAV. Tuy nhiên, trong các trường hợp nhất định, capsit thành công đã được tạo ra mà không cần gây thêm đột biến.

Biến thể AAV được bộc lộ trong bản mô tả này được tạo ra ít nhất một phần qua việc sử dụng phương pháp tiến hóa trực hệ *in vivo*, như kỹ thuật được mô tả trên đây, liên quan đến việc sử dụng sàng lọc cơ xương và cơ tim của linh trưởng sau khi dùng trong tĩnh mạch. Như vậy, biến thể capsit AAV được đề cập trong bản mô tả này bao gồm một hoặc nhiều biến đổi về trình tự axit amin khiến cho việc tải nạp vào tế bào cơ của linh trưởng hiệu quả hơn so với protein capsit AAV mẹ tương ứng. Như được sử dụng trong bản mô tả này, "protein capsit AAV mẹ tương ứng" dùng để chỉ protein capsit AAV của cùng kiểu huyết thanh AAV kiểu dại hoặc AAV biến thể như

protein capsit AAV biến thể đối tượng nhưng không bao gồm một hoặc nhiều biến đổi trình tự axit amin của protein capsit AAV biến thể đối tượng. Theo các phương án cụ thể, AAV chứa protein capsit AAV biến thể như được mô tả ở đây có tính hướng toàn thân đến nhóm nhiều cơ xương và/hoặc cơ tim xuyên suốt cơ thể sau khi cung cấp hướng đích đến mô hoặc cung cấp toàn thân.

Theo một số phương án, protein capsit AAV biến thể đối tượng bao gồm peptit khác loại nằm trong khoảng từ 5 axit amin đến khoảng 20 axit amin được chèn bởi liên kết cộng hóa trị vào vòng GH protein capsit AAV, hoặc vòng IV, so với protein capsit AAV mẹ tương ứng. Thuật ngữ "vòng GH", hoặc vòng IV, của protein capsit AAV nghĩa là phần dễ tiếp cận dung môi mà trong lĩnh vực kỹ thuật này được gọi là vòng GH, hoặc vòng IV, của protein capsit AAV. Đối với vòng GH/vòng IV của capsit AAV, xem, ví dụ, van Vliet et al. (2006) *Mol. Ther.* 14:809; Padron et al. (2005) *J. Virol.* 79:5047; và Shen et al. (2007) *Mol. Ther.* 15:1955. Do đó, ví dụ, vị trí chèn vào có thể nằm trong khoảng axit amin 411-650 của protein capsit AAV VP1. Ví dụ, vị trí chèn vào có thể nằm trong khoảng axit amin 571-612 của AAV1 VP1, trong axit amin 570-611 của AAV2 VP1, trong axit amin 571-612 của AAV3A VP1, trong axit amin 571-612 của AAV3B VP1, trong axit amin 569-610 của AAV4 VP1, trong axit amin 560-601 của AAV5 VP1, trong axit amin 571 đến 612 của AAV6 VP1, trong axit amin 572 đến 613 của AAV7 VP1, trong axit amin 573 đến 614 của AAV8 VP1, trong axit amin 571 đến 612 của AAV9 VP1, hoặc trong axit amin 573 đến 614 của AAV10 VP1, hoặc axit amin tương ứng của biến thể bất kỳ của chúng. Người có hiểu biết trong lĩnh vực kỹ thuật này cần hiểu rằng, trên cơ sở so sánh các trình tự axit amin của protein capsit của các kiểu huyết thanh AAV khác nhau, trong đó vị trí chèn vào "tương ứng với axit amin của AAV2" sẽ là protein capsit của kiểu huyết thanh AAV bất kỳ đã cho. Xem Fig. 6 về việc sắp thẳng hàng trình tự của các SEQ ID NO:1-11 AAV kiểu dại cái mà cung cấp các vị trí axit amin ở giữa và ngang các kiểu huyết thanh AAV1, AAV2, AAV3A, AAV3B, và AAV4-10 kiểu dại (có trong tự nhiên).

Theo một số phương án, vị trí chèn vào là chỉ một vị trí chèn vào duy nhất giữa hai axit amin liền kề nằm giữa axit amin 570-614 trong VP1 của kiểu huyết thanh AAV kiểu dại hoặc biến thể AAV bất kỳ, ví dụ, vị trí chèn vào nằm giữa hai axit amin liền kề nằm trong axit amin 570-610, axit amin 580-600, axit amin 570-575, axit amin

575-580, axit amin 580-585, axit amin 585-590, axit amin 590-600, hoặc axit amin 600-614, của VP1 của kiểu huyết thanh hoặc biến thể AAV bất kỳ. Ví dụ, vị trí chèn vào có thể nằm giữa axit amin 580 và 581, axit amin 581 và 582, axit amin 583 và 584, axit amin 584 và 585, axit amin 585 và 586, axit amin 586 và 587, axit amin 587 và 588, axit amin 588 và 589, hoặc axit amin 589 và 590. Vị trí chèn vào có thể nằm giữa axit amin 575 và 576, axit amin 576 và 577, axit amin 577 và 578, axit amin 578 và 579, hoặc axit amin 579 và 580. Vị trí chèn vào có thể nằm giữa axit amin 590 và 591, axit amin 591 và 592, axit amin 592 và 593, axit amin 593 và 594, axit amin 594 và 595, axit amin 595 và 596, axit amin 596 và 597, axit amin 597 và 598, axit amin 598 và 599, hoặc axit amin 599 và 600. Ví dụ, vị trí chèn vào có thể nằm giữa axit amin 587 và 588 của AAV2, giữa axit amin 590 và 591 của AAV1, giữa axit amin 588 và 589 của AAV3A, giữa axit amin 588 và 589 của AAV3B, giữa axit amin 584 và 585 của AAV4, giữa axit amin 575 và 576 của AAV5, giữa axit amin 590 và 591 của AAV6, giữa axit amin 589 và 590 của AAV7, giữa axit amin 590 và 591 của AAV8, giữa axit amin 588 và 589 của AAV9, hoặc giữa axit amin 588 và 589 của AAV10.

Theo một số phương án, đoạn peptit chèn vào được đề cập trong bản mô tả này có chiều dài là 5 axit amin, 6 axit amin, 7 axit amin, 8 axit amin, 9 axit amin, 10 axit amin, 11 axit amin, 12 axit amin, 13 axit amin, 14 axit amin, 15 axit amin, 16 axit amin, 17 axit amin, 18 axit amin, 19 axit amin, hoặc 20 axit amin. Theo phương án khác, đoạn peptit chèn vào được đề cập trong bản mô tả này bao gồm từ 1 đến 4 axit amin đệm ở đầu amino (đầu N) và/hoặc ở đầu cacboxyl (đầu C) của đoạn peptit chèn vào bất kỳ trong số các đoạn peptit chèn vào được đề cập trong bản mô tả này. Axit amin đệm minh họa bao gồm, nhưng không chỉ giới hạn ở, leuxin (L), alanin (A), glyxin (G), serin (S), threonin (T), và prolin (P). Theo một số phương án, đoạn peptit chèn vào gồm 2 axit amin đệm ở đầu N và 2 axit amin đệm ở đầu C. Theo các phương án khác, đoạn peptit chèn vào bao gồm 2 axit amin đệm ở đầu N và 1 axit amin đệm ở đầu C.

Đoạn peptit chèn vào được đề cập trong bản mô tả này chưa được mô tả trước đó và/hoặc được chèn vào capsit AAV. Không bị giới hạn bởi bất kỳ lý thuyết nào, sự có mặt của đoạn peptit chèn vào bất kỳ trong số các đoạn peptit chèn vào được đề cập có thể tác động làm giảm ái lực của capsit biến thể đối với heparin sulfat, cái mà có

khả năng thay đổi các bước nội bào và ngoại bào trong con đường tải nạp virut. Ngoài ra, môtip thêm peptit vào được đề cập trong bản mô tả này có thể khiến cho việc tải nạp của tế bào cơ (ví dụ, tế bào cơ tim) tăng lên thông qua việc bổ sung thêm miền liên kết thụ thể bề mặt tế bào.

Theo một số các phương án được ưu tiên, đoạn peptit chèn vào bao gồm trình tự axit amin có một trong số các công thức bất kỳ dưới đây.

Theo một số khía cạnh, peptit chèn vào có thể là peptit có chiều dài từ 7 đến 10 axit amin, có công thức 1a:

$Y_1Y_2X_1X_2X_3X_4X_5X_6X_7Y_3$

trong đó mỗi Y_1-Y_3, nếu có mặt, được chọn độc lập từ Ala, Leu, Gly, Ser, Thr, Pro

X_1 được chọn từ Ala,Asn, Thr, Gly, Ser, Ala, Gln, và Asp

X_2 được chọn từ Lys, Asn, Thr, Ser, Ala, và Gln

X_3 được chọn từ Ile, Thr, Lys, Leu, Val, Asn, Asp, và Arg

X_4 được chọn từ Gln, Thr, Ile, Lys, Val, Ser, và Tyr

X_5 được chọn từ Arg, Asn, Gly, Lys, Leu, Thr, Ala, Ser, và Gln

X_6 được chọn từ Thr, Lys, Val, Gly, Ser, Ala, Arg, và Pro

X_7 được chọn từ Asp, Thr, Asn, Ile, Ala, và Ser.

Theo các khía cạnh nhất định, peptit chèn vào có công thức 1a chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25) và ASDSTKA (SEQ ID NO:26). Theo một số phương án khác, peptit chèn vào có công thức 1a không chứa trình tự axit amin được chọn từ NKTTNKD (SEQ ID NO:14), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24) và NQDYTKT (SEQ ID NO:22).

Theo một số khía cạnh, peptit chèn vào có thể là peptit có chiều dài từ 7 đến 10 axit amin, có công thức 1b:

$Y_1Y_2X_1X_2X_3X_4X_5X_6X_7Y_3$

trong đó mỗi Y_1-Y_3, nếu có mặt, được chọn độc lập từ Ala, Leu, Gly, Ser, Thr, Pro

X_1 được chọn từ Thr và Asn

X_2 được chọn từ Asn và Lys

X_3 được chọn từ Lys, Ile và Thr

X_4 được chọn từ Ile, Gln, và Thr

X_5 được chọn từ Gly, Arg và Asn

X_6 được chọn từ Val, Thr và Lys

X_7 được chọn từ Thr và Asp

Theo các phương án nhất định, peptit chèn vào có công thức 1b chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14) và TNKIGVT (SEQ ID NO:15). Theo một số phương án khác, peptit chèn vào có công thức 1b không chứa trình tự axit amin được chọn từ NKTTNKD (SEQ ID NO:14).

Theo một số khía cạnh, peptit chèn vào có thể là peptit có chiều dài từ 7 đến 10 axit amin, có công thức 1c:

$Y_1Y_2X_1X_2X_3X_4X_5X_6X_7Y_3$

Trong đó mỗi Y_1-Y_3, nếu có mặt, được chọn độc lập từ Ala, Leu, Gly, Ser, Thr, Pro

X_1 được chọn từ Thr và Asn

X_2 được chọn từ Asn và Lys

X_3 được chọn từ Lys và Ile

X_4 được chọn từ Ile và Gln

X_5 được chọn từ Gly và Arg

X_6 được chọn từ Val và Thr

X_7 được chọn từ Thr và Asp

Theo các phương án nhất định, peptit chèn vào có công thức 1c chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13) và TNKIGVT (SEQ ID NO:15).

Theo một số khía cạnh, peptit chèn vào có thể là peptit có chiều dài từ 7 đến 10 axit amin, có công thức 1d:

$Y_1Y_2X_1X_2X_3X_4X_5X_6X_7Y_3$

Trong đó mỗi Y_1-Y_3, nếu có mặt, được chọn độc lập từ Ala, Leu, Gly, Ser, Thr, Pro

X_1 được chọn từ Asn và Thr

X_2 được chọn từ Asn và Lys

X_3 được chọn từ Lys và Thr

X_4 được chọn từ Ile và Thr

X_5 được chọn từ Gly, Lys và Thr

X_6 được chọn từ Lys, Arg và Val

X_7 được chọn từ Asp, Thr và Asn

Theo các phương án nhất định, peptit chèn vào có công thức 1d chứa trình tự axit amin được chọn từ TNKIGVT (SEQ ID NO:15).

Theo các phương án nhất định, peptit chèn vào chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14) và TNKIGVT (SEQ ID NO:15). Theo các phương án nhất định, peptit chèn vào chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13) và TNKIGVT (SEQ ID NO:15).

Theo một vài phương án, peptit chèn vào chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25) và ASDSTKA (SEQ ID NO:26).

Trong các phương án được ưu tiên, peptit chèn vào có từ 1 đến 3 axit amin đệm (Y_1-Y_3) tại đầu amino và/hoặc đầu cacboxyl của trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25) và ASDSTKA (SEQ ID NO:26). Trong các phương án nhất định, peptit chèn vào được chọn từ nhóm bao gồm: LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32), LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38), LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), và LAASDSTKAA (SEQ ID NO:42).

Trong một vài phương án, đối tượng protein capsit AAV biến thể không chứa bất kỳ biến đổi trình tự axit amin nào khác ngoài việc chèn peptit từ khoảng 5 axit amin đến khoảng 20 axit amin trong vòng GH, hoặc vòng IV. Ví dụ, trong một số phương án, đối tượng protein capsit AAV biến thể bao gồm chèn thêm peptit chứa trình tự axit amin được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25),ASDSTKA (SEQ ID NO:26),LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32), LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38),

LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), và LAASDSTKAA (SEQ ID NO:42), và protein capsit AAV biến thể không chứa bất kỳ sự thay thế, chèn thêm hoặc mất axit amin nào khác (tức là protein capsit AAV biến thể chứa đột biến chèn thêm đã nêu và nếu không thì giống với protein capsit AAV tương ứng). Nói cách khác, protein capsit AAV biến thể chứa đột biến chèn đã nêu giống hệt với protein capsit AAV mẹ mà peptit được chèn vào. Một ví dụ khác, protein capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25), ASDSTKA (SEQ ID NO:26), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32), LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38), LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), và LAASDSTKAA (SEQ ID NO:42), trong đó chèn thêm peptit vào giữa axit amin 587 và 588 trong VP1 của capsit AAA2, giữa axit amin 588 và 589 trong VP1 của AAV3A, AAV3B, AAV9, hoặc AAV10; giữa axit amin 589 và 590 của VP1 trong capsit AAV7; giữa axit amin 589 và 590 của VP1 trong capsit AAV7; giữa axit amin 590 và 591 của VP1 trong capsit AAV1, AAV6, hoặc AAV8, giữa axit amin 584 và 585 của VP1 trong capsit AAV4, giữa axit amin 575 và 576 của VP1 trong capsit AAV5, trong đó nếu không thì trình tự protein capsit AAV biến thể giống với trình tự protein capsit AAV mẹ tương ứng, ví dụ, trình tự trình tự bất kỳ trong số các trình tự nêu trong các SEQ ID NO:1-12.

Theo các phương án khác, protein capsit AAV biến thể, ngoài việc chứa đột biến chèn peptit, ví dụ, như đã được đề cập đến trong bản mô tả này hoặc đã được biết đến, trong vòng GH, chứa từ 1 đến 100 đột biến thay thế hoặc mất axit amin, ví dụ, từ

khoảng 1 đến 5, từ khoảng 2 đến 4, từ khoảng 2 đến 5, từ khoảng 5 đến 10, từ khoảng 10 đến 15, từ khoảng 15 đến 20, từ khoảng 20 đến 25, từ khoảng 25-50, từ khoảng 50-100 đột biến thay thế hoặc mất axit amin so với protein capsit AAV mẹ. Do đó, trong một số phương án, protein capsit biến thể chứa trình tự axit amin tương đồng 85% hoặc nhiều hơn, 90% hoặc nhiều hơn, 95% hoặc nhiều hơn, hoặc 98% hoặc nhiều hơn, ví dụ, hoặc tương đồng 99% với capsit AAV mẹ tương ứng, ví dụ, protein capsit kiểu dại nêu trong các SEQ ID NO:1-12.

Trong các phương án tiếp theo, một hoặc nhiều đột biến thay thế axit amin là tại các gốc axit amin 35, 109, 195, 213, 222, 229, 312, 319, 330, 333, 347, 363, 427, 447, 449, 453, 490, 527, 551, 581, 585, 588, 593, 606, 649, 651, 694, 698, 708, và/hoặc 735 trong protein capsit VP1 của AAV2 như được đánh số trước khi thêm peptit vào, hoặc của các gốc axit amin tương ứng của protein capsit AAV khác. Trong một số phương án, một hoặc nhiều đột biến thay thế axit amin được chọn từ nhóm bao gồm A35P, S109T, P195L, D213N, G222S, V229I, N312K, A319T, T330A, A333S, E347K, P363L, A427D, V447F, N449D, N449K, G453R, A490T, K527Q, N551S, A581T, Y585S, R588M, A593E, W606C, K649E, R651H, W694C, I698V, V708I, và L735Q của protein capsit VP1 trong AAV2 như được đánh số trước khi thêm peptit, hoặc (các) gốc axit amin tương ứng của protein capsit AAV khác.

Trong phương án được ưu tiên, protein capsit AAV biến thể được đề xuất chứa a) đột biến chèn peptit trong vòng GH của protein capsit, trong đó peptit chèn vào chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14) và TNKIGVT (SEQ ID NO:15), và b) một hoặc nhiều đột biến thay thế axit amin sau đây so với trình tự axit amin của AAV2 (SEQ ID NO:2), hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác (tức là, khác ngoài AAV2), trong đó (các) axit amin được thế không có trong tự nhiên tại các vị trí tương ứng: A35P, S109T, P195L, D213N, G222S, V229I, N312K, A319T, T330A, A333S, E347K, P363L, A427D, V447F, N449D, N449K, G453R, A490T, K527Q, N551S, A581T, Y585S, R588M, A593E, W606C, K649E, R651H, W694C, I698V, V708I, L735Q và tổ hợp của chúng. Trong một số phương án, một hoặc nhiều đột biến thay thế axit amin được chọn từ nhóm bao gồm: V708I, V708I+A593E, V708I+S109T, V708I+T330A, A35P, V708I+R588M,V708I+W606C, V708I+W694C,

I698V,N312K+N449D+N551S+I698V+L735Q,
N312K+N449D+N551S+I698V+V708I+L735Q, V708I+N449K, và V708I+G222S. Tốt hơn nếu, vị trí chèn thêm peptit, nằm giữa axit amin 587 và 588 trong capsit AAV2, nằm giữa axit amin 587 và 588 trong capsit AAV2, nằm giữa axit amin 588 và 589 trong capsit AAV3A, AAV3B, AAV9, hoặc AAV10, nằm giữa axit amin 589 và 590 trong AAV7, nằm giữa axit amin 590 và 591 trong capsit AAV1, AAV6, hoặc AAV8, nằm giữa axit amin 584 và 585 của capsit AAV4, hoặc nằm giữa axit amin 575 và 576 của capsit AAV5.

Trong phương án được ưu tiên, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin NKIQRTD (SEQ ID NO:13) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKIQRTDA (SEQ ID NO:27) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc axit amin tương ứng của capsit AAV khác, và còn chứa đột biến thay thế axit amin V708I tại vị trí 708 so với trình tự axit amin trong capsit AAV2 (SEQ ID NO:2) và tùy ý còn chứa thêm đột đột biến thay thế A593E và/hoặc S109T và/hoặc T330A và/hoặc R588M so với AAV2 hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, trong đó các axit amin được thế không có trong tự nhiên tại vị trí tương ứng. Theo một phương án cụ thể được ưu tiên khác, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin NKIQRTD (SEQ ID NO:13) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKIQRTDA (SEQ ID NO:27) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc axit amin trương ứng trong capsit AAV khác, và còn chứa thêm đột biến thay thế axit amin A35P tại gốc 35 so với trình tự axit amin trong capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác. Capsit AAV biến thể có thể tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98%, hoặc ít nhất khoảng 99%, hoặc nhiều hơn với toàn bộ chiều dài của trình tự axit amin như nêu trong ID NO:2 hoặc capsit AAV mẹ tương ứng. Trong phương án cụ thể được ưu tiên khác, capsit AAV biến thể có thể tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98%, hoặc 100% với trình tự axit amin sau đây: MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE

RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS
SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT
EQYGSVSTNLQRGN**LANKIQRTDA**RQAATADVNTQGVLPGMVWQDRDVYL
QGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFAS
FITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSINVDFTVDTNGVYS
EPRPIGTRYLTRNL (SEQ ID NO:43)

Theo một phương án cụ thể được ưu tiên khác, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin NKIQRTD (SEQ ID NO:13) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKIQRTDA (SEQ ID NO:27) nằm giữa axit amin 587 và 588 của protein capsit AAV2 hoặc axit amin trương ứng trong capsit AAV khác, và còn chứa thêm đột biến thay thế axit amin N312K so với trình tự axit amin trong capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác hoặc còn tùy ý chứa thêm (i) đột biến thay thế axit amin N449D, N551S, I698V và L735Q hoặc (ii) N449D, N551S, I698V, L735Q và V708I so với trình tự axit amin trong capsit AAV2 hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác. Capsit AAV biến thể có thể có ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98%, hoặc nhiều hơn, trình tự axit amin giống với toàn bộ chiều dài của trình tự axit amin như nêu trong ID NO:2. Trong phương án cụ thể được ưu tiên khác, capsit AAV biến thể có thể tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98%, hoặc 100% với trình tự axit amin sau đây:
MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK
YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE
RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS

SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RL**K**FKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRT**D**TPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKT**S**VDIEKVMITDEEEIRTTNPVAT
EQYGSVSTNLQRGN**LANKIQRTDA**RQAATADVNTQGVLPGMVWQDRDVYL
QGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFAS
FITQYSTGQVSVEIEWELQKENSKRWNPE**V**QYTSNYNKSVNVDFTVDTNGVY
SEPRPIGTRYLTRN**Q** (SEQ ID NO:44)

Trong phương án khác, protein capsit AAV biến thể được đề xuất chứa a) đột biến chèn peptit nằm giữa axit amin 588 và 589 của VP1 trong AAV3A, AAV3B, AAV9, hoặc AAV10, nằm giữa axit amin 589 và 590 của VP1 trong AAV7, nằm giữa axit amin 590 và 591 của VP1 trong AAV1, AAV6 hoặc AAV8, nằm giữa axit amin 584 và 585 của VP1 trong AAV4, nằm giữa axit amin 575 và 576 của VP1 trong AAV5, đột biến chèn peptit chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13) và LANKIQRTDA (SEQ ID NO:27), và b) đột biến thay thế valin thành isoleuxin tại axit amin 709 trong AAV3A hoặc AAV3B, đột biến thay thế alanin thành isoleuxin tại axit amin 709 trong AAV1 hoặc AAV6, đột biến thay thế asparagin thành isoleuxin tại axit amin 707 trong AAV4 hoặc axit amin 709 trong AAV9 hoặc đột biến thay thế threonin thành isoleuxin tại vị trí axit amin 710 trong AAV7 hoặc axit amin 711 trong AAV8 hoăc AAV10 hoặc đột biến thay thế axit amin glutamin thành isoleuxin tại axit amin 697 trong AAV5 và nếu không thì giống với trình tự bất kỳ trong số các trình tự các SEQ ID NO: 1 và 3-12. Trong phương án được ưu tiên, protein capsit biến thể chứa a) đột biến chèn peptit chứa trình tự axit amin NKIQRTD (SEQ ID NO:13) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKIQRTDA (SEQ ID NO:27) nằm giữa axit amin 587 và 588 trong capsit AAV2 và b) đột biến thay thế axit amin valin thành isoleuxin tại axit amin 708 so với trình tự axit amin trong AAV2, trong đó protein capsit biến thể có từ 2 đến 5, từ 5 đến 10, hoặc

từ 10 đến 15 đột biến thay thế axit amin.

Theo phương án khác, protein capsit biến thể chứa a) đột biến chèn peptit chứa trình tự axit amin NKIQRTD (SEQ ID NO:13) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKIQRTDA (SEQ ID NO:27) nằm giữa axit amin 587 và 588 trong capsit AAV2 và b) đột biến thay thế axit amin valin thành isoleuxin tại axit amin 708 so với trình tự axit amin trong AAV2, và nếu không thì giống với trình tự axit amin trong trình tự SEQ ID NO:2.

Trong phương án khác nữa, protein capsit biến thể chứa a) đột biến chèn peptit chứa trình tự axit amin NKIQRTD (SEQ ID NO:13) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKIQRTDA (SEQ ID NO:27) nằm giữa axit amin 587 và 588 trong capsit AAV2 và nếu không thì giống với trình tự axit amin trong trình tự SEQ ID NO:2. Trong một số phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLP
GYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADA
EFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPV
EPDSSSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTM
ATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWG
FRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAH
QGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFT
FSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQ
AGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDS
LVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTN
PVATEQYGSVSTNLQRGN**LANKIQRTDA**RQAATADVNTQGVLPGMVWQDR
DVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSA
AKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVDT
NGVYSEPRPIGTRYLTRNL (SEQ ID NO:45).

Trong phương án cụ thể được ưu tiên khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin TNKIGVT (SEQ ID NO:15) hoặc chứa, gần như

chỉ chứa, hoặc chỉ chứa trình tự axit amin LATNKIGVTA (SEQ ID NO:29) hoặc LATNKIGVTS (SEQ ID NO:37) nằm giữa axit amin 587 và 588 của capsit AAV2 hoặc vị trí tương ứng trong protein capsit của kiểu huyết thanh AAV khác và chứa đột biến thay thế axit amin V708I so với trình tự axit amin của AAV2 hoặc đột biến thay thế tương ứng của kiểu huyết thanh mẹ AAV khác và còn tùy ý chứa đột biến N449K và/hoặc G222S so với AAV2 hoặc đột biến thay thế tương ứng trong protein capsit của kiểu huyết thanh mẹ AAV, trong đó axit amin được thế không có trong tự nhiên tại vị trí tương ứng. Trong phương án được ưu tiên khác, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin TNKIGVT (SEQ ID NO:15) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LATNKIGVTA (SEQ ID NO:29) hoặc LATNKIGVTS (SEQ ID NO:37) nằ giữa axit amin 587 và 588 của capsit AAV2 hoặc vị trí tương ứng trong protein capsit của kiểu huyết thanh AAV khác và chứa N312K, N449D, N551S, I698V và L735Q và tùy ý đột biến thay thế axit amin V708I so với trình tự axit amin của AAV2 hoặc đột biến thay thế axit amin tương ứng trong kiểu huyết thanh mẹ AAV, trong đó axit amin được thế không có trong tự nhiên tại vị trí tương ứng. Capsit AAV có thể tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc lớn hơn, với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2. Trong phương án cụ thể được ưu tiên khác, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK
YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE
RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS
SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT

EQYGSVSTNLQRGN**LATNKIGVTA**RQAATADVNTQGVLPGMVWQDRDVYL QGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFAS FITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSINVDFTVDTNGVYS EPRPIGTRYLTRNL (SEQ ID NO:46)

Trong phương án khác, protein capsit AAV biến thể được đề xuất chứa a) đột biến chèn peptit nằm giữa axit amin 588 và 589 của VP1 trong AAV3A, AAV3B, AAV9, hoặc AAV10, nằm giữa axit amin 589 và 590 của VP1 trong AAV7, nằm giữa axit amin 590 và 591 của VP1 trong AAV1, AAV6 hoặc AAV8, nằm giữa axit amin 584 và 585 của VP1 trong AAV4, nằm giữa axit amin 575 và 576 của VP1 trong AAV5, đột biến chèn peptit chứa trình tự axit amin được chọn từ TNKIGVT (SEQ ID NO:15), LATNKIGVTA (SEQ ID NO:29) và LATNKIGVTS (SEQ ID NO:37), và b) đột biến thay thế valin thành isoleucin tại axit amin 709 trong AAV3A hoặc AAV3B, đột biến thay thế asparagin thành isoleuxin tại axit amin 707 trong AAV4 hoặc axit amin 709 của AAV9 hoặc đột biến thay thế threonin thành isoleuxin tại vị trí axit amin 710 trong AAV7 hoặc axit amin 711 trong AAV8 hoăc AAV10 hoặc đột biến thay thế axit amin glutamin thành isoleuxin tại axit amin 697 của AAV5. Trong phương án được ưu tiên, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin TNKIGVT (SEQ ID NO:15) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LATNKIGVTA (SEQ ID NO:29) hoặc LATNKIGVTS (SEQ ID NO:37) nằm giữa axit amin 587 và 588 trong capsit AAV2 và chứa đột biến thay thế axit amin valin thành isoleuxin tại axit amin 708 (V708I) so với trình tự axit amin trong AAV2, trong đó protein capsit biến thể có từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin và tốt hơn nếu tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc lớn hơn, với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2.

Trong phương án khác, protein capsit biến thể được đề xuất chứa a) đột biến chèn peptit chứa trình tự axit amin TNKIGVT (SEQ ID NO:15) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LATNKIGVTA (SEQ ID NO:29) hoặc LATNKIGVTS (SEQ ID NO:37) nằm giữa axit amin 587 và 588 trong capsit AAV2 và b) đột biến thay thế axit amin valin thành isoleuxin tại vị trí axit amin 708 so với trình tự axit amin AAV2 và nếu không thì giống với trình tự axit amin SEQ ID NO:2.

Trong phương án khác nữa, protein capsit biến thể được đề xuất chứa a) đột biến chèn peptit chứa trình tự axit amin TNKIGVT (SEQ ID NO:15) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LATNKIGVTA (SEQ ID NO:29) hoặc LATNKIGVTS (SEQ ID NO:37) nằm giữa axit amin 587 và 588 trong capsit AAV2 và b) đột biến thay thế axit amin valin thành isoleuxin tại vị trí axit amin 708 so với trình tự axit amin AAV2 và nếu không thì giống với trình tự axit amin SEQ ID NO:2. Theo một số phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK
YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE
RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS
SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT
EQYGSVSTNLQRGN**LATNKIGVTA**RQAATADVNTQGVLPGMVWQDRDVYL
QGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFAS
FITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVDTNGVYS
EPRPIGTRYLTRNL (SEQ ID NO:47).

Trong phương án được ưu tiên, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin NKTTNKD (SEQ ID NO:14) hoặc LANKTTNKDA (SEQ ID NO:28) nằm giữa axit amin 587 và 588 của capsit AAV2 và còn chứa thêm đột biến thay thế axit amin V708I tại vị trí 708 so với trình tự axit amin AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác và còn tùy ý chứa thêm đột biến thay thế axit amin S109T và/hoặc W694C và/hoặc W606C so với trình tự axit amin AAV2 hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, trong đó axit amin được thế không có trong tự nhiên tại vị trí tương ứng.

Trong phương án cụ thể được ưu tiên khác, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin NKTTNKD (SEQ ID NO:14) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKTTNKDA (SEQ ID NO:28) nằm giữa axit amin 587 và 588 của capsit AAV2 hoặc axit amin tương ứng trong capsit AAV khác, và còn chứa đột biến thay thế axit amin I698V tại vị trí 698 so với trình tự axit amin AAV2 (SEQ ID NO:2) hoặc gốc tương ứng trong capsit AAV khác. Capsit AAV biến thể có thể đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc nhiều hơn so với toàn bộ chiều dài của trình tự axit amin nêu trong SEQ ID NO:2 hoặc capsit AAV mẹ tương ứng. Theo phương án được ưu tiên cụ thể, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK
YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE
RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS
SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT
EQYGSVSTNLQRGN**LANKTTNKDA**RQAATADVNTQGVLPGMVWQDRDVY
LQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFA
SFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSINVDFTVDTNGVY
SEPRPIGTRYLTRNL (SEQ ID NO:48).

Theo phương án được cụ thể được ưu tiên, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin NKTTNKD (SEQ ID NO:14) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKTTNKDA (SEQ ID NO:28) nằm giữa axit amin 587 và 588 của protein capsit AAV2 hoặc vị trí tương ứng trong protein capsit của kiểu huyết thanh AAV khác, và chứa đột biến thay thế axit amin N312K so

với trình tự axit amin AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác và còn tùy ý chứa đột biến thay thế axit amin N449D, N551S, I698V, và L735Q và tùy ý V708I so với trình tự axit amin trong capsit AAV2 hoặc thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác. Capsit AAV biến thể có thể đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc nhiều hơn so với toàn bộ chiều dài của trình tự axit amin nêu trong SEQ ID NO:2. Theo phương án được ưu tiên cụ thể, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:
MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK
YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE
RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS
SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RL**K**FKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRT**D**TPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKT**S**VDIEKVMITDEEEIRTTNPVAT
EQYGSVSTNLQRGN**LANKTTNKDA**RQAATADVNTQGVLPGMVWQDRDVY
LQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFA
SFITQYSTGQVSVEIEWELQKENSKRWNPE**V**QYTSNYNKSVNVDFTVDTNGV
YSEPRPIGTRYLTRN**Q** (SEQ ID NO:49)

Trong phương án khác, protein capsit AAV biến thể được đề xuất chứa a) đột biến chèn peptit nằm giữa axit amin 588 và 589 của VP1 trong AAV3A, AAV3B, AAV9, hoặc AAV10, nằm giữa axit amin 589 và 590 trong AAV7, nằm giữa axit amin 590 và 591 trong AAV1, AAV6 hoặc AAV8, nằm giữa axit amin 584 và 585 trong AAV4, hoặc nằm giữa axit amin 575 và 576 trong AAV5, đột biến chèn peptit chứa trình tự axit amin được chọn từ NKTTNKD (SEQ ID NO:14) và LANKTTNKDA (SEQ ID NO:28), và b) đột biến thay thế valin thành isoleuxin tại axit amin 709 của AAV3A hoặc AAV3B, đột biến thay thế alanin thành isoleuxin tại

axit amin 709 của AAV1 hoặc AAV6, đột biến thay thế asparagin thành isoleuxin tại vị trí axit amin 707 của AAV4 hoặc axit amin 709 của AAV9 hoặc đột biến thay thế threonin thành isoleuxin tại vị trí axit amin 710 của AAV7 hoặc axit amin 711 của AAV8 hoặc AAV10 hoặc đột biến thay thế axit amin glutamin thành isoleuxin tại axit amin 697 của AAV5 và nếu không thì tùy ý giống với trình tự bất kỳ trong số các trình tự nêu trong các SEQ ID NO: 1 và 3-12. Theo phương án được ưu tiên, protein capsit biến thể chứa a) đột biến chèn peptit chứa trình tự axit amin NKTTNKD (SEQ ID NO:14) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKTTNKDA (SEQ ID NO: 28) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và b) đột biến thay thế axit amin valin thành isoleuxin tại axit amin 708 so với trình tự axit amin AAV2, trong đó protein capsit biến thể chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin.

Theo phương án khác, protein capsit biến thể chứa a) đột biến chèn peptit chứa trình tự axit amin NKTTNKD (SEQ ID NO:14) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKTTNKDA (SEQ ID NO:28) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và b) đột biến thay thế axit amin valin thành isoleuxin tại axit amin 708 so với trình tự axit amin AAV2, và nếu không thì giống với trình tự nêu trong SEQ ID NO:2.

Theo phương án khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin NKTTNKD (SEQ ID NO:14) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKTTNKDA (SEQ ID NO:28) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVL
PGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHAD
AEFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSP
VEPDSSSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNT
MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTY
NNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNW
GFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSA

HQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNF
TFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFS
QAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRD
SLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTT
NPVATEQYGSVSTNLQRGN**LANKTTNKDA**RQAATADVNTQGVLPGMVWQ
DRDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTF
SAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTV
DTNGVYSEPRPIGTRYLTRNL (SEQ ID NO:50).

Theo các phương án khác, protein capsit AAV biến thể được đề xuất chúa a) đột biến chèn peptit vào trong vòng GH trong protein capsit, trong đó đột biến chèn peptit chứa trình tự axit amin được chọn từ GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25) và ASDSTKA (SEQ ID NO:26), và b) một hoặc nhiều đột biến thay thế axit amin sau đây so với trình tự axit amin của AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác (tức là, ngoài AAV2), trong đó đột biến thay thế axit amin không có trong tự nhiên tại các vị trí tương ứng: A35P, S109T, P195L, D213N, G222S, V229I, N312K, A319T, T330A, A333S, E347K, P363L, A427D, V447F, N449D, N449K, G453R, A490T, K527Q, N551S, A581T, Y585S, R588M, A593E, W606C, K649E, R651H, W694C, I698V, V708I, L735Q và tổ hợp của chúng. Theo một số phương án, một hoặc nhiều đột biến thay thế axit amin được chọn từ nhóm bao gồm: V708I, S109T, R651H, A319T, P195L, P363L, I698V, D213N, G453R và tổ hợp của chúng. Theo một số phương án được ưu tiên, một hoặc nhiều đột biến thay thế axit amin bao gồm ít nhất đột biến thay thế V708I and/or P363L hoặc đột biến thay thế tương ứng tại kiểu huyết thanh mẹ AAV khác. Tốt hơn nếu, vị trí chèn thêm peptit nằm ở giữa axit amin 587 và 588 trong capsit AAV hoặc vị trí tương ứng trong protein capsit trong kiểu huyết thanh AAV khác.

Theo một số phương án, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin GNLTKGN (SEQ ID NO:16) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAGNLTKGNA (SEQ ID NO:30) nằm giữa axit amin 587 và

588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit của kiểu huyết thanh AAV khác, và còn chứa một hoặc nhiều đột biến thay thế axit amin sau đây so với trình tự axit amin trong capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương tứng trong kiểu huyết thanh mẹ AAV khác, trong đó axit amin được thế không có trong tự nhiên tại các vị trí tương ứng: V708I, V708I+S109T, R651H, A319T+P195L, P363L, P363L+V708I. Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin GNLTKGN (SEQ ID NO:16) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAGNLTKGNA (SEQ ID NO:30) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc gốc tương ứng trong capsit AAV khác và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo các phương án khác, capsit AAV biến thể chứa (i) đột biến chèn chứa trình tự axit amin GNLTKGN (SEQ ID NO:16) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAGNLTKGNA (SEQ ID NO:30) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc axit amin tương ứng trong capsit AAV khác, và (ii) chứa đột biến thay thế axit amin P363L so với trình tự axit amin AAV2 (SEQ ID NO:2) hoặc gốc tương ứng trong capsit AAV khác và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin GNLTKGN (SEQ ID NO:16) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAGNLTKGNA (SEQ ID NO:30) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc axit amin tương ứng trong capsit AAV khác, và (ii) chứa đột biến thay thế axit amin R651H so với trình tự axit amin AAV2 (SEQ ID NO:2) hoặc gốc tương ứng trong trình tự axit amin capsit AAV khác và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin GNLTKGN (SEQ ID NO:16) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAGNLTKGNA (SEQ ID NO:30) nằm giữa axit

amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:
MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK
YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE
RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS
SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT
EQYGSVSTNLQRGN**LAGNLTKGNA**RQAATADVNTQGVLPGMVWQDRDVY
LQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFA
SFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVDTNGVY
SEPRPIGTRYLTRNL (SEQ ID NO:51)

Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin NTVKLST (SEQ ID NO:17) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANTVKLSTA (SEQ ID NO:31) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyế thanh mẹ AAV khác, trong đó axit amin được thế không có trong tự nhiên tại vị trí tương ứng, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin NTVKLST (SEQ ID NO:17) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANTVKLSTA (SEQ ID NO:31) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu

trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây: MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT EQYGSVSTNLQRGN**LANTVKLSTA**RQAATADVNTQGVLPGMVWQDRDVYL QGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFAS FITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVDTNGVYS EPRPIGTRYLTRNL (SEQ ID NO:52).

Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin SNTVKAI (SEQ ID NO:18) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LASNTVKAIA (SEQ ID NO:32) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin SNTVKAI (SEQ ID NO:18) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LASNTVKAIA (SEQ ID NO:32) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%,

ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây: MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT EQYGSVSTNLQRGN**LANTVKLSTA**RQAATADVNTQGVLPGMVWQDRDVYL QGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFAS FITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVDTNGVYS EPRPIGTRYLTRNL (SEQ ID NO:53).

Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin ASNITKA (SEQ ID NO:19) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAASNITKAA (SEQ ID NO:33) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin ASNITKA (SEQ ID NO:19) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAASNITKAA (SEQ ID NO:33) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây: MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK

YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE
RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS
SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT
EQYGSVSTNLQRGN**LAASNITKAA**RQAATADVNTQGVLPGMVWQDRDVYL
QGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFAS
FITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVDTNGVYS
EPRPIGTRYLTRNL (SEQ ID NO:54).

Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin DNTVTRS (SEQ ID NO:20) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LADNTVTRSA (SEQ ID NO:34) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo một số phương án khác, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin DNTVTRS (SEQ ID NO:20) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LADNTVTRSA (SEQ ID NO:34) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin I698V so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyế thanh mẹ AAV khác, trong đó axit amin được thế không có trong tự nhiên tại vịt rí tương ứng, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, protein capsit biến

thể chứa đột biến chèn peptit chứa trình tự axit amin DNTVTRS (SEQ ID NO:20) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LADNTVTRSA (SEQ ID NO:34) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK
YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE
RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS
SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT
EQYGSVSTNLQRGN**LADNTVTRSA**RQAATADVNTQGVLPGMVWQDRDVYL
QGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFAS
FITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVDTNGVYS
EPRPIGTRYLTRNL (SEQ ID NO:55).

Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin NKISAKD (SEQ ID NO:21) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKISAKDA (SEQ ID NO:35) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin NKISAKD (SEQ ID NO:21) hoặc chứa, gần như

chỉ chứa, hoặc chỉ chứa trình tự axit amin LANKISAKDA (SEQ ID NO:35) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây: MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT EQYGSVSTNLQRGN**LANKISAKDA**RQAATADVNTQGVLPGMVWQDRDVYL QGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFAS FITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVDTNGVYS EPRPIGTRYLTRNL (SEQ ID NO:56).

Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin NQDYTKT (SEQ ID NO:22) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANQDYTKTA (SEQ ID NO:36) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo một số phương án khác, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin NQDYTKT (SEQ ID NO:22) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANQDYTKTA (SEQ ID NO:36) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong

protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin I698V so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyế thanh mẹ AAV khác (tức là, ngoài AAV2), trong đó axit amin được thế không có trong tự nhiên tại vịt rí tương ứng, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin NQDYTKT (SEQ ID NO:22) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LANQDYTKTA (SEQ ID NO:36) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:
MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVLPGYK
YLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE
RLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVEPDS
SSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNTMATGS
GAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLY
KQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPK
RLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCL
PPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTF
EDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFSQAGAS
DIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVNP
GPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVAT
EQYGSVSTNLQRGN**LANQDYTKTA**RQAATADVNTQGVLPGMVWQDRDVY
LQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFA
SFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVDTNGVY
SEPRPIGTRYLTRNL (SEQ ID NO:57).

Theo một số phương án, capsit AAV biến thể chứa đột biến chèn peptit chứa trình tự axit amin QADTTKN (SEQ ID NO:23) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAQADTTKNA (SEQ ID NO:39) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit của kiểu huyết

thanh AAV khác, và còn chứa một hoặc nhiều đột biến thay thế axit amin sau đây so với trình tự axit amin trong capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương tứng trong kiểu huyết thanh mẹ AAV khác, trong đó axit amin được thế không có trong tự nhiên tại các vị trí tương ứng: V708I, D213N, P363L, G453R. Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin QADTTKN (SEQ ID NO:23) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAQADTTKNA (SEQ ID NO:39) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo các phương án khác, capsit AAV biến thể chứa (i) đột biến chèn chứa trình tự axit amin QADTTKN (SEQ ID NO:23) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAQADTTKNA (SEQ ID NO:39) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc axit amin tương ứng trong capsit AAV khác, và (ii) chứa đột biến thay thế axit amin P363L so với trình tự axit amin AAV2 (SEQ ID NO:2) hoặc gốc tương ứng trong capsit AAV khác và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin QADTTKN (SEQ ID NO:23) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAQADTTKNA (SEQ ID NO:39) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc axit amin tương ứng trong capsit AAV khác, và (ii) chứa đột biến thay thế axit amin D213N so với trình tự axit amin AAV2 (SEQ ID NO:2) hoặc gốc tương ứng trong trình tự axit amin capsit AAV khác và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin QADTTKN (SEQ ID NO:23) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAQADTTKNA (SEQ ID NO:39) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii)

chứa đột biến thay thế axit amin G453R so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin QADTTKN (SEQ ID NO:23) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAQADTTKNA (SEQ ID NO:39) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVL
PGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHAD
AEFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSP
VEPDSSSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNT
MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTY
NNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNW
GFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSA
HQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNF
TFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFS
QAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRD
SLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTT
NPVATEQYGSVSTNLQRGN**LAQADTTKNA**RQAATADVNTQGVLPGMVWQ
DRDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTF
SAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTV
DTNGVYSEPRPIGTRYLTRNL (SEQ ID NO:58)

Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin TNRTSPD (SEQ ID NO:24) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LATNRTSPDA (SEQ ID NO:40) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác, và

chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVL
PGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHAD
AEFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSP
VEPDSSSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNT
MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTY
NNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNW
GFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSA
HQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNF
TFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFS
QAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRD
SLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTT
NPVATEQYGSVSTNLQRGN**LATNRTSPDA**RQAATADVNTQGVLPGMVWQD
RDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFS
AAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSINVDFTVD
TNGVYSEPRPIGTRYLTRNL (SEQ ID NO:59)

Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin SNTTQKT (SEQ ID NO:25) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LASNTTQKTA (SEQ ID NO:41) nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong protein capsit AAV và nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVL
PGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHAD
AEFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSP
VEPDSSSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNT

MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTY
NNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNW
GFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSA
HQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNF
TFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFS
QAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRD
SLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTT
NPVATEQYGSVSTNLQRGN**LASNTTQKTA**RQAATADVNTQGVLPGMVWQD
RDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFS
AAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVD
TNGVYSEPRPIGTRYLTRNL (SEQ ID NO:60)

Theo một số phương án, capsit AAV biến thể chứa (i) đột biến chèn peptit chứa trình tự axit amin ASDSTKA (SEQ ID NO:26) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAASDSTKAA (SEQ ID NO:42) nằm giữa axit amin 587 và 588 của CP1 trong AAV2 hoặc vị trí tương ứng trong protein capsit AAV khác, và (ii) chứa đột biến thay thế axit amin V708I so với trình tự axit amin capsit AAV2 (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyế thanh mẹ AAV khác, trong đó axit amin được thế không có trong tự nhiên tại vịt rí tương ứng, và chứa từ 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng. Theo phương án khác, protein capsit biến thể chứa đột biến chèn peptit chứa trình tự axit amin ASDSTKA (SEQ ID NO:26) hoặc chứa, gần như chỉ chứa, hoặc chỉ chứa trình tự axit amin LAASDSTKAA (SEQ ID NO:42) nằm giữa axit amin 587 và 588 của protein capsit AAV2 và nếu không thì giống với trình tự nêu được nêu trong SEQ ID NO:2. Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVL
PGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHAD
AEFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSP
VEPDSSSGTGKAGQQPARKRLNFGQTGDADSVPDPQPLGQPPAAPSGLGTNT
MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTY

NNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNW
GFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSA
HQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNF
TFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTNTPSGTTTQSRLQFS
QAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRD
SLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTT
NPVATEQYGSVSTNLQRGN**LAASDSTKAA**RQAATADVNTQGVLPGMVWQD
RDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFS
AAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKSVNVDFTVD
TNGVYSEPRPIGTRYLTRNL (SEQ ID NO:61).

Theo một vài khía cạnh khác, protein capsit AAV biến thể được đề xuất chứa một hoặc nhiều đột biến thay thế axit amin so với protein capsit AAV mẹ tương ứng, trong đó protein capsit biến thể, khi có mặt trong virion AAV khiến cho khả năng lây nhiễm vào tế bào cơ (tế bào cơ tim hoặc cơ xương) tăng lên so với khả năng lây nhiễm vào tế bào cơ bởi virion AAV chứa protein capsit AAV mẹ tương ứng.

Theo một số phương án, protein capsit AAV biến thể chứa đột biến thay thế axit amin tại axit amin 363 so với trình tự axit amin của capsit AAV2 (SEQ ID NO:2) hoặc kiểu huyết thanh mẹ AAV khác (tức là, khác AAV2). Theo một số phương án ưu tiên, protein capsit biến thể chứa trình tự axit amin đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% hoặc lớn hơn, với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 và chứa đột biến thay thế axit amin tại axit amin 363 so với trình tự axit amin của capsit AAV2 (SEQ ID NO:2). Theo một số phương án được ưu tiên, protein capsit AAV biến thể chứa đột biến thay thế axit amin P363L so với trình tự axit amin của capsit AAV2 (SEQ ID NO:2), trình tự axit amin của capsit AAV3 (SEQ ID NO:3), hoặc trình tự axit amin của capsit AAV3B (SEQ ID NO:4) hoặc đột biến thay thế axit amin P364L so với trình tự axit amin của capsit AAV1 (SEQ ID NO:1), trình tự axit amin của capsit AAV6 (SEQ ID NO: 7); hoặc đột biến thay thế axit amin P354L so với trình tự axit amin của capsit AAV4 (SEQ ID NO:5), trình tự axit amin của capsit AAV5 (SEQ ID NO:6); hoặc đột biến thay thế axit amin P365L so với trình tự axit amin của capsit AAV7 (SEQ ID NO:8), trình tự axit amin của capsit AAV9 (SEQ ID NO:10); giăhc đột biến thay thế axit amin P366L so với trình tự axit amin của capsit AAV8 (SEQ ID

NO:9), trình tự axit amin của capsit AAV10 (SEQ ID NO:11). Theo một số phương án được ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin P363L so với trình tự axit amin nêu trong (SEQ ID NO:2), hoặc đột biến thay thế tương ứng so với trình tự axit amin nêu trong SEQ ID NO: 1 và 3-12, và tương đồng ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98%, ít nhất khoảng 99% hoặc nhiều hơn so với toàn bộ chiều dài của trình tự axit amin nêu trong SEQ ID NO:2, hoặc trình tự bất kỳ nêu trong các SEQ ID NO: 1 và 3-12. Theo một số phương án được ưu tiên, protein capsit biến thể chứa trình tự axit amin chứa đột biến thay thế axit amin P363L so với trình tự axit amin nêu trong (SEQ ID NO:2) và nếu trong thì giống với trình tự axit amin nêu trong (SEQ ID NO:2). Theo một số phương án, protein capsit biến thể chứa đột biến thay thế axit amin P363L so với trình tự axit amin nêu trong (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong kiểu huyết thanh mẹ AAV khác (tức là, ngoài AAV2) trong đó protein capsit biến thể chứa từ 1 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin so với trình tự axit amin trong protein capsit AAV2 nêu trong SEQ ID NO:2 hoặc so với trình tự axit amin của protein capsit trong kiểu huyết thanh mẹ AAV khác. Theo phương án được ưu tiên khác, protein capsit biến thể chứa đột biến thay thế axit amin P363L và còn chứa đột biến thay thế axit amin E347K and/or V708I so với trình tự axit amin nêu trong SEQ ID NO:2 hoặc đột biến thay thế tương ứng trong capsit từ kiểu huyết thanh mẹ AAV khác (tức là, ngoài AAV2). Theo phương án được ưu tiên khác, protein capsit biến thể chứa đột biến thay thế axit amin P363L so với trình tự axit amin nêu trong (SEQ ID NO:2) hoặc đột biến thay thế tương ứng trong capsit từ kiểu huyết thanh mẹ AAV khác và còn chứa thêm đột biến chèn peptit, tốt hơn nếu nằm giữa axit amin 587 và 588 trong VP1 của AAV2, axit amin 588 và 589 trong AAV3A, AAV3B, AAV9, hoặc AAV10 AAV2, axit amin 589 và 590 của VP1 trong AAV7, axit amin 584 và 585 của VP1 trong AAV4, hoặc axit amin 575 và 576 trong AAV5, trong đó đột biến chèn peptit tốt hơn nếu chứa trình tự axit amin được chọn từ NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25), ASDSTKA (SEQ ID NO:26), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA

(SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32), LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38), LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), và LAASDSTKAA (SEQ ID NO:42), tốt hơn nữa nếu được chọn từ GNLTKGN (SEQ ID NO:16), LAGNLTKGNA (SEQ ID NO:30), QADTTKN (SEQ ID NO:23) và LAQADTTKNA (SEQ ID NO:39), và tùy ý chứa 2 đến 5, từ 5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin hoặc nếu không thì giống với trình tự axit amin SEQ ID NO:2 hoặc trình tự protein capsit AAV mẹ tương ứng.

Theo một số phương án khác, protein capsit AAV biến thể chứa đột biến thay thế axit amin tại axit amin 593 so với trình tự axit amin của capsit AAV2 (SEQ ID NO:2) hoặc kiểu huyết thanh mẹ AAV khác (tức là, khác AAV2). Theo một số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin tại axit amin 593 so với trình tự axit amin trong capsit AAV2 (SEQ ID NO:2) và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% hoặc lớn hơn, với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì giống với trình tự axit amin nêu trong AAV2 (SEQ ID NO: 2. Theo một số phương án, protein capsit biến thể chứa đột biến thay thế axit amin glycine thành glutamate tại axit amin 594 so với trình tự axit amin trong AAV1, AAV3A, AAV6, hoặc AAV9, hoặc tại axit amin 583 trong AAV5, hoặc tại axit amin 596 trong AAV8 hoặc AAV10, hoặc đột biến thay thế axit amin arginine thành glutamate tại axit amin 594 trong AAV3B, hoặc đột biến thay thế axit amin aspartate thành glutamate tại axit amin 592 trong AAV4 hoặc đột biến thay thế axit amin glutamin thành glutamate tại vị trí 595 trong AAV7. Theo một số phương án, protein capsit biến thể chứa đột biến thay thế axit amin A593E so với trình tự axit amin trong AAV2 và không chứa một hoặc nhiều đột biến thay thế axit amin sau đây so với trình tự axit amin trong AAV2: I19V, V369A, K26R, N215D, G355S, V46A và S196P. Theo một số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin tại axit amin A593E và V708I so với trình tự axit amin trong capsit AAV2 và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất

khoảng 98% hoặc ít nhất khoảng 99% với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì giống với trình tự axit amin nêu trong AAV2 (SEQ ID NO: 2. Theo một số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin tại axit amin A593E và S109T so với trình tự axit amin trong capsit AAV2 và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì giống với trình tự axit amin nêu trong AAV2 (SEQ ID NO: 2. Theo một số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin tại axit amin A593E, V708I và S109T so với trình tự axit amin trong capsit AAV2 và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì giống với trình tự axit amin nêu trong AAV2 (SEQ ID NO: 2. Theo một số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin tại axit amin A593E, V708I và N551S so với trình tự axit amin trong capsit AAV2 và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì giống với trình tự axit amin nêu trong AAV2 (SEQ ID NO: 2. Theo một số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin tại axit amin A593E, V708I và K649E so với trình tự axit amin trong capsit AAV2 và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì giống với trình tự axit amin nêu trong AAV2 (SEQ ID NO: 2. Theo một số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin tại axit amin A593E, V708I, S109T và K527Q so với trình tự axit amin trong capsit AAV2 và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì giống với trình tự axit amin nêu trong AAV2 (SEQ ID NO: 2.

Theo một số phương án, protein capsit AAV biến thể chứa đột biến thay thế axit amin tại axit amin 708 so với trình tự axit amin của capsit AAV2 (SEQ ID NO:2) hoặc kiểu huyết thanh mẹ AAV khác (tức là, khác AAV2), trong đó axit amin được thế

không có trong tự nhiên tại vị trí tương ứng. Tốt hơn nếu, virion rAAV không chứa đột biến thay thế axit amin proline thành serine tại axit amin 250 so với AAV2 hoặc axit amin tương ứng trong kiểu huyết thanh mẹ AAV. Theo một số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin tại axit amin 708 so với trình tự axit amin trong capsit AAV2 (SEQ ID NO:2) và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% hoặc lớn hơn, với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO: 2 Theo số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin valin thành isoleuxin (V708I) tại vị trí axit amin 708 so với trình tự axit amin trong capsit AAV2 và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% hoặc lớn hơn, với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì giống với trình tự axit amin nêu trong SEQ ID NO: 2, trong đó protein capsit biến thể không chứa đột biến thay thế axit amin P250S. Theo một số phương án, protein capsit biến thể chứa đột biến thay thế axit amin valin thành isoleuxin tại axit amin 709 trong AAV3A hoặc AAV3B, đột biến thay thế axit amin alanin thành isoleuxin tại vị trí 709 trong AAV1 hoặc AAV6, đột biến thay thế axit amin asparagin thành isoleuxin tại axit amin 707 trong AAV4 hoặc axit amin 709 trong AAV9 hoặc đột biến thay thế axit amin threonin thành isoleuxin tại axit amin 710 trong AAV7 hoặc axit amin 711 trong AAV8 hoặc AAV10 hoặc đột biến thay thế axit amin glutamin thành isoleuxin tại axit amin 697 trong AAV5. Trong phương án liên quan, protein capsit biến thể chứa đột biến thay thế axit amin V708I so với trình tự axit amin trong AAV2, trong đó protein capsit biến thể chứa từ 2 đến 5, từ5 đến 10, hoặc từ 10 đến 15 đột biến thay thế axit amin và trong đó protein capsit biến thể không chứa đột biến thay thế axit amin P250S. Theo một số phương án, protein capsit biến thể chứa đột biến thay thế axit amin V708I và cũng chứa đột biến thay thế axit amin A333S and/or S721L so với trình tự axit amin trong AAV2: Theo một số phương án ưu tiên, protein capsit biến thể chứa đột biến thay thế axit amin tại axit amin V708I, và cũng chứa đột biến thay thế axit amin A333S and/or S721L so với trình tự axit amin trong capsit AAV2 và tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% hoặc ít nhất khoảng 99% với toàn bộ chiều dài trình tự axit amin như nêu trong SEQ ID NO:2 hoặc nếu không thì

giống với trình tự axit amin nêu trong SEQ ID NO: 2.

Theo một số phương án khác, protein capsit AAV biến thể tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với trình tự capsit AAV kiểu dại được chọn từ nhóm bao gồm các SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 10,11 và 12 và cũng chứa i) một hoặc nhiều đột biến thay thế axit amin được chọn từ nhóm bao gồm A35P, D213N, A319T+P195L, P363L, P363L+V708I, G453R, R651H, I698V, V708I, V708I+A593, V708I+S109T, V708I+T330A, V708I+R588M, V708I+W694C, V708I+W606C, V708I+N449K, V708I+G222S, N312K+N449D+N551S+I698V+L735Q, N312K+N449D+N551S+I698V+V708I+L735Q, và/hoặc (ii) đột biến chèn peptit được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25), ASDSTKA (SEQ ID NO:26), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32), LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38), LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), vàLAASDSTKAA (SEQ ID NO:42). Theo một số phương án, protein capsit AAV biến thể chứa một hoặc nhiều đột biến thay thé axit amin và/hoặc đột biến chèn peptit và nếu không thì giống với trình tựu được chọn từ nhóm bao gồm các SEQ ID NO: 1-12.

Trong một số phương án, protein capsit AAV biến thể là protein capsit tổ tiên bao gồm một hoặc nhiều đột biến chèn peptit và / hoặc đột biến thay thế axit amin như được mô tả trong tài liệu này. Theo một protein capsit của tổ tiên, nó có nghĩa là tổ tiên tiến hóa của protein capsit được tìm thấy trong tự nhiên ngày nay, ví dụ: AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAVrh10, AAV11, AAV12, AAV13, được tạo ra trong silico do đột biến thay thế axit amin ngẫu nhiên ở

các vị trí của protein AAV.

Trong các phương án khác, protein capsit AAV biến thể là một loại thể khảm bao gồm các axit amin 130-725 của capsit AAV5 (SEQ ID NO: 6) hoặc trình tự axit amin tương đồng ít nhất 90%, ít nhất 95% hoặc ít nhất 98% với nó.

Trong một số khía cạnh, protein capsit AAV biến thể là thể khảm bao gồm (i) axit amin 1-129 của AAV6 (ID SEQ NO: 7) hoặc trình tự axit amin tương đồng ít nhất 90%, ít nhất 95% hoặc ít nhất 98% với nó và (ii) axit amin 130-725 của AAV5 (SEQ ID NO: 6) hoặc trình tự axit amin tương đồng ít nhất 90%, ít nhất 95% hoặc ít nhất 98% với nó và bao gồm đột biến thay thế axit amin V229I, A490T và A581T và tùy ý V447F hoặc Y585S so với trình tự của AAV5 (ID SEQ NO: 6). Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKANQQKQDDGRGLV
LPGYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHA
DAEFQERLQEDTSFGGNLGRAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDH
FPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTMSAGGGGP
LGDNNQGADGVGNASGDWHCDSTWMGDRIVTKSTRTWVLPSYNNHQYREI
KSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFRPRS
LRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPYVVGNGTEGCLP
AFPPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNF
EEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGVQFNKNLAGRYANT
YKNWFPGPMGRTQGWNLGSGVNRASVSAFTTTNRMELEGASYQVPPQPNG
MTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAY
NVGGQMATNNQSSTTAPTTGTYNLQEIVPGSVWMERDVYLQGPIWAKIPETG
AHFHPSPAMGGFGLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTV
EMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLT
RPL (SEQ ID NO:62)

Trong một số khía cạnh khác, protein capsit AAV biến thể là thể khảm bao gồm (i) axit amin 1-61 của AAV2 (ID SEQ NO: 2) hoặc trình tự axit amin tương đồng ít nhất 90%, ít nhất 95% hoặc ít nhất 98% với nó,(ii) axit amin 62-129 của AAV6 (SEQ ID NO: 7) hoặc trình tự axit amin tương đồng ít nhất 90%, ít nhất 95% hoặc ít nhất

98% với nó, và (iii) axit amin 130-725 của AAV5 (SEQ ID NO:6) và còn có thêm đột biến thay thế axit amin V229I, A490T và A581T và tùy ý V447F hoặc Y585S so với đến trình tự của AAV5 (ID SEQ NO: 6). Theo một vài phương án, capsit AAV biến thể có trình tự axit amin tương đồng ít nhất khoảng 85%, ít nhất khoảng 90%, ít nhất khoảng 95%, ít nhất khoảng 98% với hoặc giống 100% với trình tự axit amin sau đây:

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDSRGLVL
PGYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHAD
AEFQERLQEDTSFGGNLGRAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDHFP
KRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTMSAGGGGPLG
DNNQGADGVGNASGDWHCDSTWMGDRIVTKSTRTWVLPSYNNHQYREIKS
GSVDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFRPRSLR
VKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPYVVGNGTEGCLPAF
PPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEE
VPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGGVQFNKNLAGRYANTYK
NWFPGPMGRTQGWNLGSGVNRASVSAFTTTNRMELEGASYQVPPQPNGMTN
NLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVG
GQMATNNQSSTTAPTTGTYNLQEIVPGSVWMERDVYLQGPIWAKIPETGAHF
HPSPAMGGFGLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEME
WELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL
(SEQ ID NO:63).

Biến thể AAV được đề cập trong bản mô tả này được tạo ra nhờ việc sử dụng kỹ thuật tiến hóa trực hệ *in vivo* liên quan đến việc sử dụng sàng lọc tế bào cơ xương và cơ tim của linh trưởng sau khi dùng theo đường tĩnh mạch. Theo một số phương án, protein capsit biến thể được đề cập trong bản mô tả này, khi có mặt trong virion AAV, khiến cho sự tải nạp của tế bào cơ tăng lên so với sự tải nạp của tế bào cơ bởi virion AAV chứa protein capsit AAV mẹ hoặc AAV kiểu dại tương ứng. Ví dụ, theo một số phương án, protein capsit biến thể được đề cập trong bản mô tả này, khi có mặt trong virion AAV, khiến cho sự tải nạp của tế bào cơ của linh trưởng hiệu quả hơn so với virion AAV chứa protein capsit AAV mẹ hoặc protein capsit AAV kiểu dại tương ứng, ví dụ, tế bào cơ hấp thụ nhiều virion AAV chứa protein capsit AAV biến thể hơn so với virion AAV chứa protein capsit AAV mẹ hoặc AAV kiểu dại. Theo một số phương án, virion AAV biến thể hoặc rAAV biến thể có sự tải nạp của tế bào cơ tăng

lên ít nhất 2 lần, ít nhất 5 lần, ít nhất 10 lần, ít nhất 15 lần, ít nhất 20 lần, ít nhất 25 lần, ít nhất 50 lần, hoặc nhiều hơn 50 lần, so với sự tải nạp của tế bào cơ bởi virion AAV kiểu dại hoặc rAAV chứa protein capsit AAV mẹ tương ứng. Theo các phương án được ưu tiên, virion AAV biến thể hoặc rAAV biến thể có sự tải nạp của tế bào cơ gia tăng ít nhất 2 lần, ít nhất 5 lần, ít nhất 10 lần, ít nhất 15 lần, ít nhất 20 lần, ít nhất 25 lần, ít nhất 50 lần, ít nhất 100 lần, ít nhất 1000, hoặc hơn 1000 lần so với sự tải nạp của tế bào cơ bởi virion AAV8 hoặc AAV9 kiểu dại. Theo các phương án nhất định, protein capsit biến thể được đề cập trong bản mô tả này, khi có mặt trong virion AAV, khiến cho sự tải nạp rộng hơn trong tế bào cơ linh trưởng so với virion AAV chứa protein capsit AAV mẹ tương ứng hoặc protein capsit AAV kiểu dại. Nói cách khác, virion AAV biến thể tải nạp loại tế bào không được tải nạp bởi virion bao gồm protein capsit AAV mẹ tương ứng, và do đó tải nạp nhiều loại tế bào cơ hơn so với virion AAV mẹ tương ứng. Theo một số phương án, virion AAV biến thể tải nạp ưu tiên tế bào cơ, ví dụ, virion rAAV đối tượng lây nhiễm tế bào cơ với độ đặc hiệu cao hơn 2 lần, 5 lần, 10 lần, 15 lần, 20 lần, 25 lần, 50 lần, hoặc hơn 50 lần, so với tế bào cơ khác hoặc tế bào không phải tế bào cơ. Theo một số phương án, tế bào cơ đã được tải nạp là tế bào cơ tim (ví dụ, tế bào cơ tim, nguyên bào sợi tim). Theo một số phương án, tế bào cơ là tế bào cơ xương (ví dụ, nguyên bào cơ, ống cơ hoặc tế bào vệ tinh). Việc tăng sự tải nạp của tế bào cơ, ví dụ, hiệu quả tải nạp tăng lên, tải nạp rộng hơn, tải nạp được ưu tiên hơn, v..v. có thể được đánh giá một cách dễ dàng *in vitro* hoặc *in vivo* bằng phương pháp bất kỳ trong lĩnh vực kỹ thuật này để đo sự biểu hiện của gen. Ví dụ, AAV có thể được đóng gói bằng hệ gen bao gồm băng biểu hiện chứa gen báo cáo, ví dụ, protein huỳnh quang, dưới sự kiểm soát của vùng gen khởi đầu thông thường hoặc đặc hiệu mô, và mức độ tải nạp được đánh giá bằng cách phát hiện protein huỳnh quang bằng, ví dụ, kính hiển vi huỳnh quang. Đối với ví dụ khác, AAV có thể được đóng gói bằng hệ gen bao gồm trình tự axit nucleic mã vạch, và mức độ tải nạp được đánh giá bằng cách phát hiện trình tự axit nucleic bằng, ví dụ, PCR. Đối với ví dụ khác, AAV có thể được đóng gói bằng hệ gen bao gồm băng biểu hiện chứa gen trị liệu để điều trị bệnh cơ, và mức độ tải nạp được đánh giá bằng cách phát hiện sự điều trị bệnh cơ ở bệnh nhân mắc bệnh mà được dùng AAV.

Bệnh có thể được điều trị bằng cách sử dụng vectơ rAAV hoặc virion biến thể và/hoặc phương pháp được đề cập trong bản mô tả này bao gồm, nhưng không chỉ giới

hạn ở, bệnh do gen đơn, bệnh phức tạp, và các tổn thương do chấn thương. Ví dụ về bệnh do gen đơn bao gồm, nhưng không chỉ giới hạn ở, loạn dưỡng cơ chẳng hạn như loạn dưỡng cơ Duchenne, Becker, bẩm sinh (bao gồm, nhưng không chỉ giới hạn ở bệnh cơ Bethlem, loạn dưỡng cơ Ullrich, loạn dưỡng cơ Fukuyama, thiếu hụt Intfrin, loạn dưỡng cơ thiếu hụt merosin và hội chứng Walker-Warburgh), ngoại biên (bao gồm, nhưng không chỉ giới hạn ở Gowers-Laing, Miyoshi, và Nonaka), loạn dưỡng cơ Emery-Dreifuss, loạn dưỡng cơ thể mặt-vai-cánh tay, loạn dưỡng cơ vùng gốc chi, loạn dưỡng cơ tăng trương lực cơ và loạn dưỡng cơ; rối loạn cơ bắp bẩm sinh và chứng giật cơ bẩm sinh, bệnh cơ myotubular, bệnh cơ lõi trung tâm; bệnh sợi cơ, bệnh sợi cơ liên quan đến desmin, bệnh thiếu máu, hội chứng Andersen-Tawil, bệnh cơ Nemaline, bệnh Brody, rối loạn dự trữ lysosome, chẳng hạn như thiếu hụt alpha-mannosidaza trong lysosome, aspartylglucosaminuria, thiếu hụt beta-mannosidaza trong lysosome, xystein niêu, bệnh Farber, fucosidosis, bệnh Gaucher, thiếu hụt galactosidaza, gangliosidoses (bao gồm, nhưng không chỉ giới hạn ở biến thể AB, thiếu hụt chất hoạt hóa, thiếu hụt beta-galactosidaza, bệnh Fabry, bệnh Sandhoff, và bệnh Schindler), rối loạn dự trữ glycogen (bao gồm, nhưng không chỉ giới hạn ở bệnh Andersen, bệnh Cori, bệnh Danon, bệnh Forbes, thiếu hụt glucozo-6-phosphat, bệnh Hers, thiếu hụt lactat dehydrogenaza A, bệnh Pompe, bệnh Tarui, và bệnh von Gierke), bệnh dự trữ axit sialic tự do ở trẻ em, thiếu hụt axit lipaza tiêu thể, bệnh Krabbe, loạn dưỡng chất trắng nhược sắc, bệnh mucopolysaccarit (bao gồm, nhưng không chỉ giới hạn ở thiếu hụt hyaluronidaza, hội chứng Hunter, hội chứng Hurler, hội chứng Hurler-Scheie, hội chứng Maroteaux-Lamy, hội chứng Morquio, hội chứng Sanfilippo, hội chứng Scheie, và hội chứng Sly), bệnh mucolipidosis (bao gồm, nhưng không chỉ giới hạn ở Sialidosis, bệnh I-cell, thiếu hụt mucolipidin 1, và Psuedy-Hurler Polydystrophy), thiếu hụt đa sulfataza, bệnh Niemann-Pick, bệnh lý tích tụ lipofuscin ở neuron (bao gồm, nhưng không chỉ giới hạn ở bệnh Batten-Spielmeyer-Vogt, thiếu hụt Cathepsin D bẩm sinh, bệnh Batten dạng muộn ở trẻ em Đức/Xéc-bi, bệnh Jansky-Bielschowsky, bệnh Kufs, bệnh Batten dạng muộn ở trẻ em, bệnh Batten dạng muộn ở trẻ em, bệnh động kinh Northern, bệnh Santavuori-Haltia, và Turkish Late Infantile), xơ cứng xương, bệnh Salla, thiếu hụt Saposin B, bệnh Tay-Sach và bệnh Wolman; rối loạn chuyển hóa chẳng hạn như thiếu hụt adenosin monophosphat deaminaza, Alcapton niệu, thiếu hụt carnitin, thiếu hụt carnitin palmityl transferaza, rối loạn

Hartnup, Homocystin niệu, bệnh siro niệu, thiếu hụt myophosphorylaza, thiếu hụt phosphofuctokinaza, thiếu hụt phosphoglycerat kinaza, thiếu hụt phosphoglycerat mutaza, thiếu hụt phosphorylaza, và bệnh Tangier; mất điều hòa Friedreich; mất điều hòa talengiectasia; mất điều hòa với thiếu thụt vitamin E; liệt chu kỳ, chẳng hạn như bệnh Gamstorp và liệt chu kỳ kali máu; bệnh ti thể chẳng hạn như hội chứng Barth, hội chứng Kearns-Sayre, bệnh cơ ti thể, bệnh não ty lạp thể, nhiễm axit lactic, và các cơn đột quỵ, động kinh giật cơ với sợi cơ không đều, và hội chứng Pearson; bệnh cơ tim phì đại di truyền; bệnh cơ tim giãn nở; bệnh tim bẩm sinh di tuyền, chẳng hạn như bệnh van động mạch chủ di truyền và không co bóp của tâm thất trái có khuyết tật tim bẩm sinh; rối loạn nhịp tim có yếu tố gia đình, chẳng hạn như liệt chu kỳ rối loạn nhịp tim Anderson, khuyết tật vách ngăn tâm nhĩ với khuyết tật dẫn truyền AV, hội chứng Brugada, khuyết tật dẫn tuyền tim, nhịp tim nhanh thất đa hình liên quan đến catecholamin, và chẹn tim bẩm sinh; rối loạn mạch máu di truyền, chẳng hạn như hội chứng động mạch tortuosity, bệnh động mạch não di truyền thể trội gây nhồi máu dưới màng cứng và bệnh não chất trắng, bệnh động mạch não di truyền thể lặn gây nhồi máu dưới màng cứng và bệnh não chất trắng, phình động mạch chủ bụng có yếu tố gia đình, hội chứng Marfan, hội chứng Ehlers-Danlos, hội chứng Beals bẩm sinh, hội chứng Loeys-Dietz, và giả u vàng sợi chun; bệnh cơ tim thất phải gây loạn nhịp; bệnh cơ tim thất phải gây loạn nhịp có yếu tố gia đình; bệnh Naxos; không co bóp tâm thất trái; bệnh cơ tâm nhĩ có yếu tố gia đình; chứng nhịp tâm thất đập nhanh có yếu tố gia đình; hội chứng Wolff-Parkinson-White di truyền; hội chứng QT dài; hội chứng QT ngắn; hội chứng xoang bệnh lý; bệnh lipoprotein, chẳng hạn như chứng thiếu betalipoprotein và thiếu hụt lipoprotein lipaza; thiếu hụt alpha-1 antitrypsin; thiếu hụt yếu tố đông máu VIII (hemophilia A) thiếu hụt hoặc yếu tố đông máu IX (hemophilia B); bệnh thiếu tan máu bẩm sinh; chứng vôi hóa các cơ khớp; bệnh lamila dạng lão hóa; bệnh Huntington; hội chứng nhược cơ bẩm sinh; hội chứng Hutchinson-Gilford Progeria; hội chứng Noonan; bệnh cơ thiếu cân đối dạng sợi bẩm sinh; xơ hóa bẩm sinh cơ ngoài mắt; bệnh cơ cấu trúc lõi nhỏ; bệnh cơ gợn sóng; hội chứng Schwartz-Jampel; bệnh cơ tổng hợp hình ống; và bệnh cơ thể vằn. Ví dụ về bệnh phức tạp bao gồm, nhưng không chỉ giới hạn ở, bệnh tim/tim mạch (ví dụ, suy tim sung huyết, nhồi máu cơ tim, chứng đau thắt ngực, bệnh động mạch vành, bệnh tim thiếu máu cục bộ, bệnh cơ tim); ung thư; tiểu đường; và lây nhiễm. Ví dụ về tổn thương do chấn thương,

nhưng không chỉ giới hạn ở, nhiễm virut trong cơ, rách cơ; và giập cơ. Theo các phương án được ưu tiên, vectơr rAAV hoặc virion biến thể và/hoặc phương pháp được mô tả trong bản môt tả này được sử dụng để điều trị bệnh Fabry, mất điều hòa Friedreich's, loạn dưỡng cơ Duchenne, loạn dưỡng cơ Becker, bệnh Pompe, thiếu hụt myophosphorylaza, loạn dưỡng cơ thể mặt-vai-cánh tay, loạn dưỡng cơ vòng đai gốc chi, hoặc loạn dưỡng cơ tăng trương lực cơ.

Theo phương án khác, capsit biến thể được đề cập trong bản mô tả này bao gồm axit nucleic khác loại chứa trình tự nucleotit mã hóa sản phẩm gen như, nhưng không chỉ giới hạn ở, ARN can thiệp, ARN không mã hóa dài, ARN không mã hóa ngắn, ARN đối nghĩa, aptame, polypeptit, kháng thể được tiết ra, kháng thể đơn chuỗi, miền V_{HH}, thụ thể hòa tan, affibody, knottin, DARPin, xenturin, chaperon, nucleaza đặc hiệu vị trí tạo ra đặc hiệu vị trí sự bất hoạt quá trình phiên mã chức năng của gen hoặc nucleaza đặc hiệu vị trí được biến đổi tạo ra sự hoạt hóa đặc hiệu gen của quá trình phiên mã.

Virion rAAV biến thể được đề cập trong bản mô tả này bao gồm axit nucleic khác loại chứa trình tự nucleotit mã hóa sản phẩm gen. Theo một số phương án, sản phẩm gen là ARN đối nghĩa, microRNA (miRNA), ARN kẹp tóc ngắn, (shRNA), ARN can thiệp nhỏ hoặc tiền chất hoặc sản phẩm tương tự của chúng. Theo một số phương án, sản phẩm gen là ARN không mã hóa dài. Theo một số phương án, sản phẩm gen là ARN không mã hóa ngắn. Theo một số phương án, sản phẩm gen là ARN đối nghĩa. Theo một số phương án, sản phẩm gen là aptame. Theo một số phương án, sản phẩm gen là polypeptit. Theo một số phương án, sản phẩm gen là kháng thể được tiết ra. Theo một số phương án, sản phẩm gen là kháng thể đơn chuỗi. Theo một số phương án, sản phẩm gen là miền V_{HH}. Theo một số phương án, sản phẩm gen là thụ thể hòa tan. Theo một số phương án, sản phẩm gen là affibody. Theo một số phương án, sản phẩm gen là knottin. Theo một số phương án, sản phẩm gen là DARPin. Theo một số phương án, sản phẩm gen là centurin. Theo một số phương án, sản phẩm gen là chaperon. Theo một số phương án, sản phẩm gen là nucleaza đặc hiệu vị trí cung cấp cho sự bất hoạt quá trình phiên mã đặc hiệu vị trí của chức năng của gen.

Việc sử dụng sản phẩm gen bao gồm, nhưng không chỉ giới hạn ở, làm tăng hàm lượng của yếu tố trong tế bào, làm tăng làm lượng của yếu tố ở tế bào lân cận qua sự tiết yếu tố, làm giảm hàm lượng của yếu tố trong tế bào, hoặc làm giảm hàm lượng

của yếu tố ở tế bào lân cận hoặc tế bào xa qua sự tiết yếu tố. Sản phẩm gen có thể được thiết kế để bổ sung hàm lượng thiếu hụt của sản phẩm gen bị mất, làm giảm hàm lượng thiếu hụt của sản phẩm gen bị mất, đưa vào sản phẩm gen hỗ trợ mới, chất bổ sung mức sản phẩm gen hỗ trợ, làm giảm mức sản phẩm gen cản trợ, hoặc cả làm giảm hàm lượng sản phẩm gen cản trợ lẫn đưa vào hoặc bổ sung hàm lượng của sản phẩm gen hỗ trợ.

Sản phẩm gen được vận chuyển bởi biến thể AAV đối tượng có thể được sử dụng để thay đổi hàm lượng của sản phẩm gen hoặc hoạt tính của sản phẩm gen liên hệ trực tiếp hoặc gián tiếp với bệnh và tổn thương cơ. Cơ xương, cơ tim hoặc cơ trơn được tải nạp với biến thể AAV đối tượng cũng được sử dụng làm yếu tố sinh học để sản xuất và tiết protein trị liệu để điều trị bệnh *in trans* cho cơ quan ở xa. Các gen mà sản phẩm gen của nó được liên hệ trực tiếp hoặc gián tiếp với các bệnh di truyền bao gồm, ví dụ, gen mã hóa cho sản phẩm gen bất kỳ trong số các sản phẩm gen sau đây: dystrophin bao gồm mini- và micro-dystrophins (DMD; ví dụ, số truy cập GenBank NP_003997.1; SEQ ID NO:64); titin (TTN); titin cap (TCAP) α-sarcoglycan (SGCA), β-sarcoglycan (SGCB), γ -sarcoglycan (SGCG) hoặc δ-sarcoglycan (SGCD); alpha-1-antitrypsin (A1-AT); chuỗi nặng myosin 6 (MYH6); chuỗi nặng myosin 7 (MYH7); chuỗi nặng myosin 11 (MYH11); chuỗi nhẹ myosin 2 (ML2); chuỗi nhẹ myosin 3 (ML3); chuỗi nhẹ myosin kinaza 2 (MYLK2); protein C liên kết với myosin (MYBPC3); desmin (DES); dynamin 2 (DNM2); laminin a2 (LAMA2); lamin A/C (LMNA); lamin B (LMNB); thụ thể lamin B (LBR); dysferlin (DYSF); emerin (EMD); insulin; yếu tố đông máu, bao gồm, nhưng không chỉ giới hạn ở, yếu tố VIII và yếu tố IX; erythropoietin (EPO); lipoprotein lipaza (LPL); lưới cơ tương $Ca2^{++}$-ATPaza (SERCA2A), protein liên kết S100 canxi A1 (S100A1); myotubularin (MTM); DM1 protein kinaza (DMPK; ví dụ, số truy cập GenBank NG_009784.1; SEQ ID NO:65); glycogen phosphorylaza L (PYGL); glycogen phosphorylaza, liên kết cơ (PYGM; ví dụ, Số truy cập GenBank NP_005600.1; SEQ ID NO:66); glycogen synthaza 1 (GYS1); glycogen synthaza 2 (GYS2); α- galactosidaza A (GLA; ví dụ, Số truy cập GenBank NP_000160.1; SEQ ID NO:67); α-N-axetylgalactosaminidaza (NAGA); axit α-glucosidaza (GAA; ví dụ, Số truy cập GenBank NP_000143.2; SEQ ID NO:68), sphingomyelinaza phosphodiesteraza 1 (SMPD1); axit lipaza tiêu thể

(LIPA); chuỗi α1 collagen loại I (COL1A1); chuỗi α2 collagen loại I (COL1A2); chuỗi α1 collagen loại III (COL3A1); chuỗi α1 collagen loại V (COL5A1); chuỗi α2 collagen loại V (COL5A2); chuỗi α1 collagen loại VI (COL6A1); chuỗi α2 collagen loại VI (COL6A2); chuỗi α3 collagen loại VI (COL6A3); procollagen-lysin 2-oxoglutarat 5-dioxygenaza (PLOD1); axit lipaza tiêu thể (LIPA); frataxin (FXN; ví dụ, số truy cập GenBank NP_000135.2; SEQ ID NO:69); myostatin (MSTN); β-N-axetyl hexosaminidaza A (HEXA); β-N-axetylhexosaminidaza B (HEXB); β-glucocerebrosidaza (GBA); adenosin monophosphat deaminaza 1 (AMPD1); β-globin (HBB); iduronidaza (IDUA); iduronat 2-sulfat (IDS); troponin 1 (TNNI3); troponin T2 (TNNT2); troponin C (TNNC1); tropomyosin 1 (TPM1); tropomyosin 3 (TPM3); N-axetyl-α-glucosaminidaza (NAGLU); N-sulfoglucosamin sulfohydrolaza (SGSH); heparan-α-glucosaminit N-axetyltransferaza (HGSNAT); integrin α 7 (IGTA7); integrin α 9 (IGTA9); glucosamin (N-axetyl)-6-sulfataza (GNS); galactosamin (N-axetyl)-6-sulfataza (GALNS); β-galactosidaza (GLB1); β-glucuronidaza (GUSB); hyaluronoglucosaminidaza 1 (HYAL1); axit ceramidaza (ASAH1); galactosylcermidaza (GALC); cathepsin A (CTSA); cathepsin D (CTSA); cathepsin K (CTSK); chất hoạt hóa GM2 gangliosit (GM2A); arylsulfataza A (ARSA); arylsulfataza B (ARSB); enzym tạo formylglyxin (SUMF1); neuraminidaza 1 (NEU1); N-axetylglucosamin-1-phosphat transferaza α (GNPTA); N-axetylglucosamin-1-phosphat transferaza β (GNPTB); N-axetylglucosamin-1-phosphat transferaza γ (GNPTG); mucolipin-1 (MCOLN1); NPC vận chuyển nội bào (NPC1); NPC vận chuyển nội bào 2 (NPC2); xeroit lipofuscinosis 5 (CLN5); xeroit lipofuscinosis 6 (CLN6); xeroit lipofuscinosis 8 (CLN8); palmitoyl protein thioesteraza 1 (PPT1); tripeptidyl peptidaza 1 (TPP1); battenin (CLN3); DNAJ họ protein sốc nhiệt 40 chi C5 (DNAJC5); Protein 8 có chứa miền thuộc siêu họ điều hành chính (MFSD8); mannosidaza α nhóm 2B chi 1 (MAN2B1); mannosidaza β (MANBA); aspartylglucosaminidaza (AGA); α-L-fucosidaza (FUCA1); cystinosin, vận chuyển cystein tiêu thể (CTNS); sialin; chất mang tan họ 2 chi 10 (SLC2A10); chất mang tan họ 17 chi 5 (SLC17A5); chất mang tan họ 6 chi 9 (SLC6A19); chất mang tan họ 22 chi 5 (SLC22A5); chất mang tan họ 37 chi 4 (SLC37A4); protein màng liên kết với tiêu thể 2 (LAMP2); kênh α dẫn truyền natri cổng điện áp tiểu phần 4 (SCN4A); kênh dẫn

truyền natri β cổng điện áp tiểu phần 4 (SCN4B); kênh αdẫn truyền natri cổng điện áp tiểu phần 5 (SCN5A); kênh α dẫn truyền natri cổng điện áp tiểu phần (SCN4A); kênh dẫn truyền natri cổng điện áp tiểu phần α1c (CACNA1C); kênh dẫn truyền natri cổng điện áp tiểu phần α1s (CACNA1S); phosphoglycerat kinaza 1 (PGK1); phosphoglycerat mutaza 2 (PGAM2); amylo-a-1,6-glucosidaza,4-a-glucanotransferaza (AGL); phân họ liên quan đến kênh dẫn truyền kali cổng điện áp ISK- chi 1 (KCNE1); phân họ liên quan đến kênh dẫn truyền kali cổng điện áp ISK- chi 2 (KCNE2); phân họ kênh dẫn truyền kali cổng điện áp J chi 2 (KCNJ2); phân họ kênh dẫn truyền kali cổng điện áp J chi 5 (KCNJ5); phân họ kênh dẫn truyền kali cổng điện áp H chi 2 (KCNH2); phân họ tương tự kênh dẫn truyền kali cổng điện áp KQT- chi 1 (KCNQ1); kênh dẫn truyền kali cổng nucleotit vòng được hóa hóa quá phân cực (HCN4); kênh clo cổng điện áp 1 (CLCN1); carnitine palmitoyltransferaza 1A (CPT1A); thụ thể ryanodin 1 (RYR1); thụ thể ryanodin 2 (RYR2); chất tích hợp cầu nối 1 (BIN1); LARGE xylosyl- và glucuronyltransferaza 1 (LARGE1); protein docking 7 (DOK7); fukutin (FKTN); protein liên quan đến fukutin (FKRP); selenoprotein N (SELENON); protein O-mannosyltransferaza 1 (POMT1); protein O-mannosyltransferaza 2 (POMT2); protein O-liên kết với mannoza N-axetylglucosaminyltransferaza 1 (POMGNT1); protein O- liên kết với mannoza N-axetylglucosaminyltransferaza 2 (POMGNT2); protein-O-mannose kinaza (POMK); miền isoprenoit synthaza chứa protein (ISPD); plectin (PLEC); tiểu phần epsilon nicotinic thụ thể tác động kiểu colin (CHRNE); colin O-axetyltransferaza (CHAT); colin kinaza β (CHKB); tiểu phần đuôi giống collagen của axetylcolinsteraza không đối xứng (COLQ); protein kết hợp thụ thể của của xynap (RAPSN); protein 1 bốn và một nửa miền LIM (FHL1); β-1,4-glucuronyltransferaza 1 (B4GAT1); β-1,3-N-axetylgalactosaminyltransferaza 2 (B3GALNT2); dystroglycan 1 (DAG1); protein xuyên màng 5 (TMEM5); protein xuyên màng 43 (TMEM43); protein 2 liên kết SECIS (SECISBP2); glucosamin (UDP-N-axetyl)-2-epimeraza/N-axetylmannosamin kinaza (GNE); anoctamin 5 (ANO5); duy trì cấu trúc của vùng bản lề linh hoạt trong nhiễm sắc thể chứa protein 1 (SMCHD1); lactat dehydrogenaza A (LDHA); lactat dehydrogenaza B (LHDB); calpain 3 (CAPN3); caveolin 3 (CAV3); mô-tip ba bên chứa protein 32 (TRIM32); protein liên kết với axit nucleic loại ngón tay kẽm CCHC- (CNBP); nebulin (NEB); actin, α1, cơ

xương (ACTA1); actin, α1, cơ tim (ACTC1); actinin α2 (ACTN2); poly(A)-liên kết protein nhân 1 (PABPN1); protein 3 chứa miền LEM 3 (LEMD3); metalloproteinaza kẽm STE24 (ZMPSTE24); protein vận chuyển triglyxerit tiêu thể(MTTP); tiểu phần α1 nicotinic thụ thể cholinergic (CHNA1); tiểu phần α2 nicotinic thụ thể cholinergic (CHRNA2); tiểu phần α3 nicotinic thụ thể cholinergic (CHRNA3); tiểu phần α4 nicotinic thụ thể cholinergic (CHRNA4); tiểu phần α5 nicotinic thụ thể cholinergic (CHRNA5); tiểu phần α6 nicotinic thụ thể cholinergic (CHRNA6); tiểu phần α7 nicotinic thụ thể cholinergic (CHRNA7); tiểu phần α8 nicotinic thụ thể cholinergic (CHRNA8); tiểu phần α9 nicotinic thụ thể cholinergic (CHRNA9); tiểu phần α10 nicotinic thụ thể cholinergic (CHRNA10); tiểu phần β1 nicotinic thụ thể cholinergic (CHRNB1); tiểu phần β2 nicotinic thụ thể cholinergic (CHRNB2); tiểu phần β3 nicotinic thụ thể cholinergic (CHRNB3); tiểu phần β4 nicotinic thụ thể cholinergic (CHRNB4); tiểu phần γ nicotinic thụ thể cholinergic (CHRNG1); tiểu phần ∂ nicotinic thụ thể cholinergic (CHRND); tiểu phần ε nicotinic thụ thể cholinergic (CHRNE1); phân họ A chi 1 băng liên kết ATP (ABCA1); phân họ C chi 6 băng liên kết ATP (ABCC6); phân họ C chi 9 băng liên kết ATP (ABCC9); phân họ D chi 1 băng liên kết ATP (ABCD1); protein 1 vận chuyển Ca2+ lưới cơ tương/lưới nội chất ATPaza (ATP2A1); ATM serin/threonin kinaza (ATM); protein tocopherol transferaza (TTPA); họ kinesin chi 21A (KIF21A); paired-like homeobox 2a (PHOX2A); heparan sulfat proteoglycan 2 (HSPG2); phân tử tương tác chất nền 1 (STIM1); notch 1 (NOTCH1); notch 3 (NOTCH3); dystrobrevin α (DTNA); protein hoạt hóa AMP kinaza, noncatalytic γ2 (PRKAG2); protein 3 giàu xystein và glyxin (CSRP3); viniculin (VCL); myozenin 2 (MyoZ2); myopalladin (MYPN); junctophilin 2 (JPH2); phospholamban (PLN); calreticulin 3 (CALR3); protein liên kết nexilin F-actin (NEXN); protein 3 liên kết miền LIM (LDB3); eyes absent 4 (EYA4); huntingtin (HTT); thụ thể androgen (AR); thụ thể phi protein tyrosin phosphat typ 11 (PTPN11); plakoglobin nối (JUP); desmoplakin (DSP); plakophilin 2 (PKP2); desmoglein 2 (DSG2); desmocollin 2 (DSC2); catenin a3 (CTNNA3); NK2 homeobox 5 (NKX2-5); protein mỏ neo A-kinaza 9 (AKAP9); protein mỏ neo A-kinaza 10 (AKAP10); polypeptit 2 có hoạt tính ức chế protein α liên kết nucleotit guanin (GNAI2); ankyrin 2 (ANK2); syntrophin a-1 (SNTA1); calmodulin 1 (CALM1); calmodulin 2 (CALM2);

HTRA serin peptidaza 1 (HTRA1); fibrillin 1 (FBN1); fibrillin 2 (FBN2); xylosyltransferaza 1 (XYLT1); xylosyltransferaza 2 (XYLT2); tafazzin (TAZ); homogentisat 1,2-dioxygenaza (HGD); tiểu phần xúc tác glucoza-6-phosphataza (G6PC); enzym 1 1,4-alpha-glucan (GBE1); phosphofructokinaza cơ (PFKM); tiểu phần alpha 1 điều hòa phosphorylaza kinaza (PHKA1); tiểu phần alpha 2 điều hòa phosphorylaza kinaza (PHKA2); tiểu phần beta điều hòa phosphorylaza kinaza (PHKB); tiểu phần gamma 2 điều hòa phosphorylaza kinaza (PHKG2); phosphoglyxerat mutaza 2 (PGAM2); xystathionin-beta-synthaza (CBS); metylentetrahydrofolat reductaza (MTHFR); 5-metyltetrahydrofolat-homoxystein metyltransferaza (MTR); 5-metyltetrahydrofolat-homoxystein metyltransferaza reductaza (MTRR); metylmalonic axituria và homocystinuria, cblD typ (MMADHC); AND ty thể, bao gồm, nhưng không chỉ giới hạn ở NADH được mã hóa trong ty thể: tiểu phần 1 lõi ubiquinon oxidoreductaza (MT-ND1); NADH được mã hóa trong ty thể: tiểu phần 5 lõi ubiquinon oxidoreductaza (MT-ND5); tRNA axit glutamic được mã hóa trong ty thể (MT-TE); tRNA histadin được mã hóa trong ty thể (MT-TH); tRNA leuxin 1 được mã hóa trong ty thể (MT-TL1); tRNA lysin được mã hóa trong ty thể (MT-TK); tRNA serin 1 được mã hóa trong ty thể (MT-TS1); tRNA valin được mã hóa trong ty thể (MT-TV); protein kinaza kinaza 1 hoạt hóa mitogen (MAP2K1); B-Raf gen đột biến gây bệnh ung thư sớm nhất, serin/threonin kinaza (BRAF); raf-1 gen đột biến gây bệnh ung thư sớm nhất, serin/threonin kinaza (RAF1); các yếu tố sinh trưởng, bao gồm, nhưng không chỉ giới hạn ở yếu tố sinh trưởng insulin 1 (IGF-1); yếu tố sinh trưởng chuyển đổi β3 (TGFβ3); thụ thế yếu tố sinh trưởng chuyển đổi β, typ I (TGFβR1); thụ thể yếu tố sinh trưởng chuyển đổi β, typ II (TGFβR2), yếu tố sinh trưởng nguyên bào sợi 2 (FGF2), yếu tố sinh trưởng nguyên bào sợi 4 (FGF4), yếu tố sinh trưởng nội mô mạch A (VEGF-A), yếu tố sinh trưởng nội mô mạch B (VEGF-B); yếu tố sinh trưởng nội mô mạch C (VEGF-C), yếu tố sinh trưởng nội mô mạch D (VEGF-D), thụ thể yếu tố sinh trưởng nội mô mạch 1 (VEGFR1), và thụ thể yếu tố sinh trưởng nội mô mạch 2 (VEGFR2); interleukin; immunoadhesin; xytokin; và kháng thể.

Theo các phương án được ưu tiên, sản phẩm gen được vận chuyển bởi các biến thể AAV đối tượng được lựa chọn từ alpha galactosidaza A (GLA), Frataxin (FXN), Dystrophin (DMD), A xitalpha glucosidaza (GAA), và Glycogen phosphorylaza, cơ

(PYGM). Theo một số phương án ưu tiên, biến thể AAV chứa đoạn axit nucleic chứa trình tự nucleotit mã hóa cho (i) polypeptit GLA bao gồm hoặc chỉ bao gồm trình tự axit amin nêu trong SEQ ID NO:67, (ii) polypeptit FXXN bao gồm hoặc chỉ bao gồm trình tự axit amin nêu trong SEQ ID NO:69, (iii) polypeptit DMD bao gồm hoặc chỉ bao gồm đoạn chức năng (ví dụ, mini hoặc micro dystrophin, tốt hơn nếu bao gồm đoạn liên kết với actin còn nguyên vẹn, ít nhất 4 trong số 24 đoạn lặp lại giống spectrin và miền liên kết dystroglycan) của trình tự axit amin nêu trong SEQ ID NO:64, (iv) polypeptit GAA bao gồm hoặc chỉ bao gồm trình tự axit amin nêu trong SEQ ID NO:68, (v) polypeptit PYGM bao gồm hoặc chỉ bao gồm trình tự axit amin nêu trong SEQ ID NO:66, (vi) hoặc (v) trình tự axit amin tương đồng ít nhất 80%, ít nhất 85%, ít nhất 90% hoặc ít nhất 95% với trình tự bất kỳ trong số trình tự nêu trong SEQ ID NO:64 và 66-69.

Theo phương án được ưu tiên khác, biến thể AAV chứa gen chuyển mã hóa cho ARN can thiệp, ví dụ, ARN đối nghĩa, miARN, shARN, hoặc si ARN, cái mà làm giảm biểu hiện của DMPK. Theo một vài khía cạnh, ARN can thiệp làm giảm biểu hiện của DMPK được mã hóa bởi axit nucleic có trình tự nucleotit như nêu trong SEQ ID NO:65 hoặc trình tự tương đồng ít nhất 80%, ít nhất 85%, ít nhất 95% với trình tự nêu trong SEQ ID NO:65.

Các gen mà sản phẩm gen của nó gây ra hoặc thúc đẩy sự gây chết tế bào theo chương trình dùng trong bản mô tả này dùng để chỉ "gen tiền gây chết tế bào theo chương trình" và sản phẩm của các gen này (mARN; protein) được gọi là "sản phẩm gen tiền gây chết tế bào theo chương trình". Đích tiền gây chết theo chương trình bao gồm, ví dụ, sản phẩm gen Bax; sản phẩm gen Bid; sản phẩm gen Bak; sản phẩm gen Bad; Bcl-2; Bcl-X1. Sản phẩm gen tiền gây chết theo chương trình bao gồm chất ức chế sự gây chết tế bào theo chương trình có liên kết với nhiễm sắc thể X.

Các gen mà sản phẩm gen của nó gây ra hoặc thúc đẩy sự hình thành mạch được dùng trong bản mô tả này dùng để chỉ "gen tiền hình thành mạch" và sản phẩm của các gen này (mARN; protein) được gọi là "sản phẩm của gen tiền hình thành mạch". Các đích tiền hình thành mạch bao gồm, ví dụ, yếu tố sinh trưởng nội mô mạch (VEGFa, VEGFb, VEGFc, VEGFd); thụ thể 1 yếu tố sinh trưởng nội mô mạch (VEGFR1); thụ thể 2 yếu tố sinh trưởng nội mô mạch (VEGFR2); Tyrosin Kinaza liên quan đến Fms 1 (Flt1); yếu tố sinh trưởng nhau thai (PGF); yếu tố sinh trưởng có

nguồn gốc tiểu cầu (PDGF); các angiopoietin; gen nhím Sonic. Các gen mà sản phẩm gen của nó ức chế sự hình thành mạch được dùng trong bản mô tả này dùng để chỉ "gen chống hình thành mạch" và sản phẩm của các gen này (mARN; protein) được gọi là "sản phẩm của gen chống hình thành mạch". Sản phẩm của gen chống hình thành mạch bao gồm endostatin; tumstatin; angiostatin; yếu tố có nguồn gốc biểu mô sắc tố (PEDF), và protein dung hợp hoặc kháng thể đặc hiệu với các đích tiền hình thành mạch và/hoặc các thụ thể của chúng, ví dụ, kháng thể đặc hiệu VEGF Avastin™, v.v..

Các gen mà có sản phẩm gen hoạt động chức năng như chất điều biến miễn dịch, ví dụ, yếu tố bổ sung, thụ thể giống Toll, được gọi là "gen điều biến miễn dịch". Gen điều biến miễn dịch minh họa bao gồm xytokin, chemokin, và protein dung hợp hoặc kháng thể đặc hiệu đối với chúng và/hoặc các thụ thể của chúng, ví dụ, protein dung hợp kháng IL-6 Rilonacept™, kháng thể đặc hiệu yếu tố bổ sung H lampamizumab, v..v. Các gen có sản phẩm gen hoạt động chức năng như yếu tố bảo vệ cơ, ví dụ, yếu tố sinh trưởng insulin 1 (IGF-1); yếu tố sinh trưởng chuyển đổi β (TGFβ); yếu tố sinh trưởng nguyên bào sợi (FGF).

Trong một số trường hợp, sản phẩm gen quan tâm là enđonucleaza đặc hiệu vị trí cái mà tạo ra để bất hoạt đặc hiệu vị trí chức năng gen, ví dụ, trong đó enđonucleaza gây bất hoạt các alen liên quan đến bệnh cơ. Ví dụ, khi alen trội mã hóa cho bản sao thiếu hụt của gen, khi mà ở dạng kiểu dại, là protein cấu trúc cơ và/hoặc tạo ra chức năng cơ bình thường, enđonucleaza đặc hiệu vị trí có thể được hướng đích đến alen thiếu hụt và gây bất hoạt alen thiếu hụt.

Ngoài việc làm bất hoạt gen alen thiếu hụt, nucleaza đặc hiệu vị trí cũng có thể được sử dụng để kích thích sự tái tổ hợp cùng loại với ADN thể cho mã hóa bản sao chức năng của protein được mã hóa bởi alen thiếu hụt. Do đó, ví dụ, virion rAAV của đối tượng có thể được sử dụng để vận chuyển cả enđonucleaza đặc hiệu vị trí cái mà gây bất hoạt alen thiếu hụt, và có thể được sử dụng để vận chuyển bản sao chức năng của alen thiếu hụt, dẫn đến việc sửa chữa alen thiếu hụt, nhờ đó tạo ra protein cơ chức năng (ví dụ, lamin A/C chức năng, fibrillin chức năng, collagen chức năng typ IV chức năng, v.v.). Theo một số phương án, virion rAAV được đề cập trong bản mô tả này bao gồm trình tự nucleotit khác loại mã hóa cho enđonucleaza đặc hiệu vị trí; và trình tự nucleotit khác loại mã hóa cho bản sao chức năng của alen thiếu hụt, trong đó bản sao chức năng mã hóa protein cơ chức năng. Protein cơ chức năng bao gồm, ví dụ,

lamin A/C, fibrillin 1, COL6A1, COL6A2, COL6A3, và tương tự.

Enđonucleaza đặc hiệu vị trí thích hợp để sử dụng bao gồm, ví dụ, meganucleaza; nucleaza ngón tay kẽm (ZFN); nucleaza hiệu ứng giống chất hoạt hóa quá trình phiên mã (TALEN); và trình tự ngắn xuôi ngược lặp lại ngăn cách bởi các khoảng trống đều đặc được tạo cụm/CRISPR kết hợp (Cas), trong đó enđonucleaza đặc hiệu vị trí này không có trong tự nhiên và được biến đổi để hướng đích đến gen đặc hiệu. Nucleaza đặc hiệu vị trí này có thể được biến đổi di truyền để cắt các vị trí đặc hiệu trong hệ gen, và đầu nối không cùng loại sau đó có thể sửa chữa phần đứt gãy trong khi chèn hoặc loại bớt một vài nucleotit. Enđonucleaza đặc hiệu vị trí này (còn được gọi là "INDEL") sau đó xê dịch protein ra khỏi khung và gây bất hoạt gen hiệu quả. Xem, ví dụ, Công bố đơn yêu cầu cấp bằng sáng chế của Mỹ số 2011/0301073.

Theo một số phương án về vectơ rAAV biến thể được đề cập trong bản mô tả này, trình tự nucleotit mã hóa sản phẩm của gen quan tâm được liên kết có điều khiển với vùng gen khởi đầu cơ bản. Vùng gen khởi đầu cơ bản thích hợp bao gồm, ví dụ, vùng gen khởi đầu virut cự bào (CMV) (Stinski et al. (1985) *Journal of Virology* 55(2): 431-441), gen khởi đầu β-actin ở gà (CBA)/chất tăng cường sớm CMV / β-globin intron thỏ (CAG) (Miyazaki et al. (1989) *Gene* 79(2): 269-277, CB^{SB} (Jacobson et al. (2006) *Molecular Therapy* 13(6): 1074-1084), vùng gen khởi đầu yếu tố 1α kéo dài ở người (EF1α) (Kim et al. (1990) *Gene* 91(2): 217-223), vùng gen khởi đầu phosphoglyxerat kinaza người (PGK) (Singer-Sam et al. (1984) *Gene* 32(3): 409-417, vùng gen khởi đầu sợi nặng ty thể (Loderio et al. (2012) *PNAS* 109(17): 6513-6518), vùng gen khởi đầu ubiquitin (Wulff et al. (1990) *FEBS Letters* 261: 101-105). Theo các phương án khác, trình tự nucleotit mã hóa sản phẩm gen quan tâm được liên kết có điều khiển với vùng gen khởi đầu cảm ứng. Trong một số trường hợp, trình tự nucleotit mã hóa sản phẩm gen quan tâm được liên kết có điều khiển với yếu tố điều hòa đặc hiệu mô hoặc đặc hiệu tế bào. Ví dụ, trong một số trường hợp, trình tự nucleotit mã hóa sản phẩm gen quan tâm được liên kết có điều khiển với yếu tố điều hòa đặc hiệu cơ (ví dụ, vùng gen khởi đầu đặc hiệu cơ tim hoặc vùng gen khởi đầu đặc hiệu cơ xương), ví dụ, yếu tố điều hòa mà tạo ra sự biểu hiện chọn lọc của gen được liên kết có điều khiển trong tế bào cơ. Yếu tố điều hòa đặc hiệu cơ thích hợp bao gồm, ví dụ, vùng gen khởi đầu α-actin cơ xương (Muscat và Kedes (1987) *Mol. Cell. Biol.* 7:4089-4099); vùng gen khởi đầu α-actin cơ tim (Minty và Kedes (1986) *Mol. Cell.*

Biol. 6:2125–2136); vùng gen khởi đầu α-actin cơ trơn (Nakano et al. (1991) *Gene* 99:285–289); vùng gen khởi đầu α-actin cơ trơn mạch máu (Keogh et al. (1999) *Gene Therapy* 6(4):616-628); vùng gen khởi đầu creatin kinaza cơ (Bartlett et al. (1996) *Cell Transplantation* 5(3):411-419); vùng gen khởi đầu chuỗi nhẹ myosin 1 và chuỗi nhẹ myosin 3 (Seidel và Arnold (1989) *J. Biol. Chem.* 264(27):16109-16117); vùng gen khởi đầu chuỗi nhẹ myosin 2v (MLC2v) (Su et al. (2004) *PNAS* 101(46):16280-16285); vùng gen khởi đầu yếu tố điều hòa myogen 5 (Myf5) (Fujimaki et al. (2004) *Journal of Biological Chemistry* 289(11):7399-7412); vùng gen khởi đầu biệt hóa myogen 1 (Myod1) (Zingg et al. (1994) *Nucleic Acids Research* 22(12):2234-2241); vùng gen khởi đầu myogen (Myog) (Salminen et al. (1991) *Journal of Cell Biology* 115(4):905-917); vùng gen khởi đầu gen hộp bắt cặp 7 (Pax7) (Murmann et al. (2000) *Biol Chem.* 381(4):331-335); vùng gen khởi đầu homeodomain giống bắt cặp 3 (Pitx3) (Coulon et al. (2007) *Journal of Biological Chemistry* 282:33192-33200); vùng gen khởi đầu MHCK7 (Salva et al. (2007) *Mol. Ther.* 15(2):320-329); vùng gen khởi đầu MCK/SV40 (Takeshita et al. (2007) *International Journal of Molecular Medicine* 19:309-315); vùng gen khởi đầu C5-12 (Li et al. (1999) *Nature Biotechnology* 17:241-245); chất tăng cường/vùng gen khởi đầu hai và ba bộ đôi trước sau MCK (Wang et al. (2008) *Gene Therapy* 15:1489-1499); vùng gen khởi đầu chuỗi nặng myosin 7 (MYH7); (Iwaki et al. (2104) *PLoS ONE* 9(4):e88610); vùng gen khởi đầu chuỗi nặng myosin 6 (MYH6) (Pacak et al. (2008) *Genet. Vaccines Ther.* 6:13); vùng gen khởi đầu chuỗi nặng troponin T tim (TNNT2) (Farza et al. (1998) *J. Mol. Cell Cardiol.* 30(6):1247-53); vùng gen khởi đầu α-tropomyosin (Helfman et al. (1986) *Molecular và Cellular Biology* 6(11):3582-3595); vùng gen khởi đầu troponin C tim (TNNC1) (Scheier et al. (1990) *Journal of Biological Chemistry* 34(5):21247-21253); vùng gen khởi đầu protein C liên kết với myosin tim (Lin et al. (2013) *PLoS ONE* 8(7):e69671); vùng gen khởi đầu troponin I tim (TNNI3) (Bhavsar et al. (1996) *Genomics* 35(1):11-23); vùng gen khởi đầu desmin (Li et al. (1991) *Journal of Biological Chemistry* 10(5):6562-6570); vùng gen khởi đầu trao đổi natri-canxi (NCX1) (Scheller et al. (1997) *Journal of Biological Chemistry* 273(13):7643-7649); vùng gen khởi đầu yếu tố natriuretic tâm nhĩ (Durocher et al. (1996) *Molecular và Cellular Biology* 16(9):4648-4655); và vùng gen khởi đầu SM22α (Kemp et al. (1995) *Biochemical Journal* 310(3):1037-1043.

Đối với mục đích của sáng chế, phần mô tả trong bản mô tả này đề xuất axit nucleic phân lập bao gồm trình tự nucleotit mã hóa capsit protein AAV biến thể như được mô tả trên đây. Axit nucleic phân lập có thể là vectơ AAV, ví dụ, vectơ AAV tái tổ hợp.

Phần mô tả trong bản mô tả này cũng đề xuất phương pháp điều trị bệnh cơ, phương pháp bao gồm cho cá thể cần điều trị dùng lượng có hiệu quả của virion rAAV biến thể chứa gen chuyển quan tâm như được mô tả trên đây và được đề cập trong bản mô tả này. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này có thể xác định một cách dễ dàng lượng có hiệu quả của virion rAAV của đối tượng và bệnh đã được điều trị bằng cách thử nghiệm sự thay đổi một hoặc nhiều thông số chức năng và giải phẫu, ví dụ, sinh thiết cơ sau đó hóa mô miễn dịch, lấy mẫu huyết thanh sau đó ELISA hoặc thử nghiệm hoạt tính enzym, thử nghiệm đi bộ, tiêu thụ oxi tối đa, phân tích sinh khối phân suất tống máu thất trái, thể tích tâm thu thất trái, đo động lực học bằng tay, nâng trọng lượng tối đa, điểm số khả năng vận động của Hammersmith, dứng dậy từ trên sàn theo thời gian hoặc thử nghiệm 9 Hole Peg.

Các phương pháp không giới hạn phạm vi của sáng chế để đánh giá chức năng cơ và các thay đổi của chúng bao gồm đánh giá thử nghiệm đi bộ, tiêu thụ oxi tối đa, phân tích sinh khối phân suất tống máu thất trái, thể tích tâm thu thất trái, thang đo Vignos, kiểm tra chức năng theo thời gian, điểm số khả năng vận động của Hammersmith, đứng lên trên sàn theo thời gian, thang đo chức năng vận động, đánh giá khả năng đi lại không phải nằm của bệnh nhân North Star, thử nghiệm 9 lỗ hoặc thử nghiệm của Bệnh viện nhi Philadelphia về rối loạn thần kinh cơ.

Theo một số phương án, lượng có hiệu quả của virion rAAV dẫn đến giảm tỷ lệ mất chức năng cơ, tính toàn vẹn cơ giải phẫu, hoặc khối lượng cơ, ví dụ, tỷ lệ mất giảm 2 lần, 3 lần, 4 lần, hoặc 5 lần hoặc nhiều hơn và do đó giảm tiến triển của bệnh, ví dụ, tỷ lệ mất giảm 10 lần hoặc nhiều hơn và do đó tiến triển của bệnh. Theo một số phương án, lượng có hiệu quả của virion rAAV dẫn đến việc thu được chức năng cơ, độ bền cơ, khối lượng cơ, và/hoặc cải thiện tính toàn vẹn cơ giải phẫu hoặc chất đánh dấu sinh học, ví dụ, cải thiện 2 lần, 3 lần, 4 lần hoặc 5 lần hoặc nhiều hơn chức năng cơ, độ bền cơ, khối lượng cơ, và/hoặc cải thiện tính toàn vẹn cơ giải phẫu hoặc chất đánh dấu sinh học, ví dụ, cải thiện 10 lần hoặc nhiều hơn chức năng cơ, độ bền cơ, khối lượng cơ, và/hoặc cải thiện tính toàn cơ giải phẫu, hoặc chất đánh dấu sinh học.

Như được đánh giá một cách dễ dàng bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, liều lượng cần để đạt được hiệu quả điều trị mong muốn sẽ thường nằm trong khoảng từ 1 x 10^8 đến khoảng 1 x 10^{16} virion tái tổ hợp, thường được đề xuất bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này là 1 x 10^8 đến khoảng 1 x 10^{16} "hệ gen vectơ" và tốt hơn nếu nằm trong khoảng từ 1 x 10^{11} đến khoảng 1 x 10^{15} virion tái tổ hợp.

Virion rAAV có thể được vận chuyển đến cơ xương bằng cách tiêm trong mạch (động mạch hoặc tĩnh mạch), tiêm trong màng bụng, truyền dịch vào chi và/hoặc tiêm trong cơ trực tiếp bằng phương thức hay đường dùng thích hợp khác dẫn đến vận chuyển virion rAAV tới cơ xương. Virion AAV có thể được vận chuyển đến cơ tim bằng cách tiêm trong mạch (động mạch hoặc tĩnh mạch, tiêm trực tiếp vào tim (vào trong tâm nhĩ trái, tâm nhĩ phải, tâm thất phải và/hoặc vách ngăn), truyền ngược dòng hoặc xuôi dòng vào trong động mạch vành (thông qua động mạch trái giảm dần hoặc động mạch vành mũ trái), tuần hoàn khép kín, tiêm trong màng ngoài tim, tiêm truyền tĩnh mạch, hoặc bằng bất kỳ cách thức hoặc con đường thuận tiện nào dẫn đến việc vận chuyển virut đến cơ tim. Trong một phương án được ưu tiên, virion rAAV được vận chuyển đến cơ xương và/hoặc cơ tim bằng cách tiêm tĩnh mạch toàn thân. Khi được tiêm tĩnh mạch, virion rAAV có thể di chuyển qua hệ thống tuần hoàn và chuyển hóa tế bào cơ một cách hiệu quả hơn, so với khả năng của virut AAV kiểu dại hoặc virion AAV chứa protein capsit AAV mẹ tương ứng.

Capsit protein biến thể được đề cập trong bản mô tả này được phân lập, ví dụ, được tinh chế. Theo một số phương án, capsit protein biến thể được đề cập trong bản mô tả này chứa trong vectơ AAV hoặc virion AAV tái tổ hợp (rAAV). Theo các phương án khác, vectơ AAV biến thể và/hoặc virion AAV biến thể như vậy được sử dụng trong phương pháp điều trị *in vivo* hoặc *ex vivo* bệnh cơ trong cơ xương hoặc cơ tim của linh trưởng.

Phần mô tả trong bản mô tả này còn đề xuất tế bào chủ như, nhưng không chỉ giới hạn ở, tế bào chủ (đã biến đổi di truyền) được phân lập bao gồm axit nucleic của đối tượng. Tế bào chủ theo sáng chế được đề cập trong bản mô tả này, có thể là tế bào phân lập, như tế bào từ môi trường nuôi cấy tế bào *in vitro*. Tế bào chủ này rất hữu ích trong việc tạo ra virion rAAV biến thể đối tượng, như được đề cập trong bản mô tả này. Theo một phương án, tế bào chủ này được biến đổi di truyền ổn định bằng axit

nucleic. Theo các phương án khác, tế bào chủ được biến đổi di truyền tạm thời bằng axit nucleic. Axit nucleic như vậy được đưa ổn định hoặc tạm thời vào tế bào chủ, bằng cách sử dụng kỹ thuật đã thiết lập bao gồm, nhưng không chỉ giới hạn ở, điện di, kết tủa canxi phosphat, lây nhiễm qua trung gian liposom, và các kỹ thuật tương tự. Để biến nạp ổn định, axit nucleic thường sẽ còn chứa thêm gen đánh dấu chọn lọc, ví dụ, gen đánh dấu bất kỳ trong số một vài gen đánh dấu chọn lọc đã biết như gen kháng neomyxin, và các gen tương tự. Tế bào chủ này được tạo ra bằng cách đưa axit nucleic vào tế bào bất kỳ trong số nhiều loại tế bào khác nhau, ví dụ, tế bào động vật có vú, bao gồm, ví dụ, tế bào chuột, và tế bào linh trưởng (ví dụ, tế bào người). Tế bào động vật có vú minh họa bao gồm, nhưng không chỉ giới hạn ở, tế bào và dòng tế bào sơ cấp, trong đó dòng tế bào minh họa bao gồm, nhưng không chỉ giới hạn ở, tế bào HEK293, tế bào HEK293T, tế bào COS, tế bào HeLa, tế bào Vero, nguyên bào sợi chuột 3T3, nguyên bào sợi C3H10T1/2, tế bào CHO, và các tế bào tương tự. Tế bào chủ minh họa bao gồm, nhưng không chỉ giới hạn ở, tế bào HeLa (ví dụ, Bộ Sưu tập môi trường nuôi cấy Mỹ (American Type Culture Collection - ATCC) số CCL-2), tế bào CHO (ví dụ, ATCC số CRL9618, CCL61, CRL9096), tế bào 293 (ví dụ, ATCC số CRL-1573), tế bào Vero, tế bào NIH 3T3 (ví dụ, ATCC số CRL-1658), tế bào Huh-7, tế bào BHK (ví dụ, ATCC số CCL10), tế bào PC12 (ATCC số CRL1721), tế bào COS, tế bào COS-7 (ATCC số CRL1651), tế bào RAT1, tế bào L của chuột (ATCC số CCLI.3), tế bào phôi thận của người (HEK) (ATCC số CRL1573), tế bào HLHepG2, và các tế bào tương tự. Tế bào chủ cũng có thể được tạo ra bằng cách sử dụng virut hình que để lây nhiễm tế bào côn trùng như tế bào Sf9, mà tạo ra AAV (xem, ví dụ, Bằng sáng chế của Mỹ số 7,271,002; Đơn yêu cầu cấp bằng sáng chế của Mỹ số 12/297,958). Theo một số phương án, tế bào chủ được biến đổi di truyền bao gồm, ngoài axit nucleic bao gồm trình tự nucleotit mã hóa capsit protein AAV biến thể, như được mô tả trên đây, axit nucleic bao gồm trình tự nucleotit mã hóa một hoặc nhiều protein Rep AAV. Theo các phương án khác, tế bào chủ còn bao gồm vectơ rAAV biến thể. Virion rAAV biến thể có thể đuọc tạo ra bằng cách sử dụng tế bào chủ này. Phương pháp tạo ra virion rAAV được mô tả trong, ví dụ, Công bố đơn yêu cầu cấp bằng sáng chế của Mỹ số 2005/0053922 và Công bố đơn yêu cầu cấp bằng sáng chế của Mỹ số 2009/0202490.

Phần mô tả trong bản mô tả này đề xuất tiếp dược phẩm chứa: a) virion rAAV

biến thể, như được mô tả trên đây và được đề cập trong bản mô tả này; và b) chất mang dược dụng, chất pha loãng, tá dược, hoặc dung dịch đệm. Theo một số phương án, chất mang dược dụng, chất pha loãng, tá dược, hoặc dung dịch đệm thích hợp để dùng ở bệnh nhân là người hoặc không phải người. Các tá dược, chất mang, chất pha loãng, và dung dịch đệm này bao gồm dược chất bất kỳ có thể dùng được mà không gây ra độc tính quá mức. Tá dược dược dụng bao gồm, nhưng không chỉ giới hạn ở, chất lỏng như nước, nước muối, glyxerol và etanol. Các muối dược dụng chứa trong đó có thể là, ví dụ, các muối axit vô cơ như hyđroclorua, hyđrobromua, phosphat, sulfat, và các muối tương tự; và các muối của các axit hữu cơ như axetat, propionat, malonat, benzoat, và các muối tương tự. Ngoài ra, chất bổ trợ, như chất thấm ướt hoặc chất nhũ hóa, chất hoạt động bề mặt, chất tạo đệm độ pH, và các chất tương tự, có thể có mặt trong các chất dẫn thuốc như vậy. Nhiều loại tá dược dược dụng khác nhau là đã biết trong lĩnh vực kỹ thuật và không cần được mô tả chi tiết trong bản mô tả này. Tá dược dược dụng đã được mô tả đầy đủ trong nhiều ấn phẩm khác nhau, bao gồm, ví dụ, A. Gennaro (2000) "Remington: The Science và Practice of Pharmacy", 20th edition, Lippincott, Williams, & Wilkins; Pharmaceutical Dosage Forms và Drug Delivery Systems (1999) H. C. Ansel et al., eds., 7th ed., Lippincott, Williams, & Wilkins; và Handbook of Pharmaceutical Excipients (2000) A. H. Kibbe et al., eds., 3rd ed. Amer. Pharmaceutical Assoc. Theo một số khía cạnh, sáng chế đề xuất dược phẩm chứa khoảng 1 x 10^8 đến khoảng 1 x 10^{16} virut tái tổ hợp hoặc 1 x 10^8 đến khoảng 1 x 10^{16} hệ gen vectơ, trong đó mỗi virut tái tổ hợp này bao gồm hệ gen mã hóa một hoặc nhiều sản phẩm gen.

Một vài phương án của sáng chế được lấy ví dụ trong mục từ 1 đến 54 sau đây:

1. Protein capsit virut adeno kết hợp (AAV) biến thể chứa đột biến chèn peptit vào vòng GH của protein capsit, trong đó đoạn chèn vào là trong AAV2 hoặc vị trí tương ứng trong phần capsit của kiểu huyết thanh AAV kiểu dại khác ngoài AAV2 hoặc biến thể AAV, và trong đó đoạn peptit chèn vào được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25), ASDSTKA (SEQ

ID NO:26), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32), LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38), LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), và LAASDSTKAA (SEQ ID NO:42).

2. AAV biến thể theo mục 1, trong đó protein capsit chứa một hoặc nhiều đột biến điểm liên quan đến AAV2 hoặc một hoặc nhiều đột biến điểm liên quan đến các kiểu huyết thanh AAV kiểu dại hoặc các biến thể AAV.

3. AAV biến thể theo mục 2, trong đó một hoặc nhiều đột biến điểm được chọn từ nhóm bao gồm A35P, S109T, P195L, D213N, G222S, V229I, N312K, A319T, T330A, A333S, E347K, P363L, A427D, V447F, N449D, N449K?, G453R, A490T, K527Q, N551S, A581T, Y585S, R588M, A593E, W606C, K649E, R651H, W694C, I698V, V708I, và L735Q, và tốt hơn nếu được chọn từ nhóm bao gồm V708I, V708I+A593E, V708I+S109T, V708I+T330A, A35P, V708I+R588M, V708I+W606C, V708I+W694C, I698V, N312K+N449D+N551S+I698V+L735Q, N312K+N449D+N551S+I698V+V708I+L735Q, V708I+N449K, và V708I+G222S.

4. AAV biến thể theo mục 1, trong đó đoạn peptit chèn vào được chèn vào sau axit amin bất kỳ trong số các axit amin ở vị trí 570-671 trong VP1 của AAV2 hoặc vị trí tương ứng trong kiểu huyết thanh AAV kiểu dại hoặc biến thể AAV khác.

5. AAV biến thể theo mục 4, trong đó đoạn peptit chèn vào được chèn vào sau axit amin ở vị trí 587 trong VP1 của AAV2 hoặc vị trí tương ứng trong kiểu huyết thanh AAV khác.

6. Virion virut adeno kết hợp tái tổ hợp (rAAV) lây nhiễm chứa: (a) protein capsit AAV biến thể theo điểm bất kỳ trong số các mục 1-5, và axit nucleic khác loại.

7. rAAV theo mục 6, trong đó axit nucleic khác loại chứa trình tự nucleotit mã hóa cho tác nhân can thiệp ARN hoặc polypeptit.

8. Phương pháp vận chuyển axit nucleic khác loại đến tế bào đích bao gồm cho tế bào đích tiếp xúc với virion rAAV theo mục 7.

9. Phương pháp theo mục 8, trong đó tế bào đích là tế bào cơ xương và/hoặc cơ

tim.

10. Phương pháp theo mục 8, trong đó tế bào đích là *in vitro*.

11. Phương pháp theo mục 8, trong đó tế bào đích là *in vivo*.

12. Axit nucleic đã phân lập chứa trình tự axit nucleic mã hóa cho protein capsit virut adeno kết hợp (AAV) biến thể chứa đột biến chèn peptit vào vòng GH của protein capsit, trong đó đoạn chèn vào là trong AAV2 hoặc vị trí tương ứng trong phần capsit trong kiểu huyết thanh AAV kiểu dại khác AAV2 hoặc biến thể AAV, và trong đó đoạn peptit chèn vào được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25), ASDSTKA (SEQ ID NO:26), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32), LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38), LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), và LAASDSTKAA (SEQ ID NO:42).

13. Tế bào chủ được phân lập chứa axit nucleic theo điểm 12.

14. Protein capsit virut adeno kết hợp (AAV) chứa đột biến chèn peptit so với protein capsit AAV mẹ chứa hai axit amin liền kề tại vị trí nằm giữa axit amin 570 và 611 trong VP1 của AAV2, trong đó đoạn thêm vào chứa trình tự axit amin $Y_1Y_2X_1X_2X_3X_4X_5X_6X_7Y_3$, và trong đó X_1 được chọn từ T và N; X_2 được chọn từ N và K; X_3 được chọn từ K, I và T; X_4 được chọn từ I, Q và T; X_5 được chọn từ G, R và N; X_6 được chọn từ V, T và K; và X_7 được chọn từ T và D.

15. AAV biến thể theo mục 4, trong đó đoạn peptit chèn vào được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14) và TNKIGVT (SEQ ID NO:15).

16. AAV biến thể theo mục 15, trong đó đột biến chèn peptit vào nằm bên sườn axit amin LA đầu N và axit amin A đầu C.

17. AAV biến thể theo mục 15, trong đó đoạn peptit chèn vào được chèn vào nằm giữa axit amin 587 và 588 trong VP1 của AAV2 hoặc vị trí tương ứng trong kiểu huyết thanh AAV hoặc AAV biến thể.

18. Virion virut adeno kết hợp tái tổ hợp (rAAV) lây nhiễm chứa (a): protein capsit AAV biến thể theo mục bất kỳ trong số các mục từ 1-17, và axit nucleic khác loại.

19. rAAV theo mục 18, trong đó axit nucleic khác loại chứa tác nhân can thiệp ARN hoặc trình tự nucleotit mã hóa polypeptit.

20. Phương pháp vận chuyển axit nucleic khác loại đến tế bào đích, bao gồm cho tế bào đích tiếp xúc với virion rAAV theo mục 18.

21. Phương pháp theo mục 20, trong đó tế bào đích là tế bào cơ xương và/hoặc cơ tim.

22. Phương pháp theo mục 21, trong đó tế bào đích là *in vitro* hoặc *in vivo*.

23. Protein capsit virut adeno kết hợp (AAV) biến thể chứa i) trình tự axit amin AAV tương đồng ít nhất 90% với AAV kiểu dại được chọn từ nhóm bao gồm các SEQ ID NO: 1-10 và 11; và ii) một hoặc nhiều đột biến thay thế axit amin được chọn từ nhóm bao gồm P363L, P363L+V708I, P363L+E347K, V708I+A593E, V708I+A333S, V708I+S721L, V708I+A593E+N551S, V708I+A593E+K649E, V708I+A593E+S109T, V708I+A593E+S109T+K527Q, A593E+S109T, trong đó một hoặc nhiều đột biến thay thế so với AAV2 hoặc một hoặc nhiều đột biến so với kiểu huyết thanh AAV khác.

24. AAV biến thể theo mục 23, trong đó protein capsit chứa đột biến chèn peptit.

25. AAV biến thể theo mục 24, trong đó đoạn peptit chèn vào được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13), NKTTNKD (SEQ ID NO:14), TNKIGVT (SEQ ID NO:15), GNLTKGN (SEQ ID NO:16), NTVKLST (SEQ ID NO:17), SNTVKAI (SEQ ID NO:18), ASNITKA (SEQ ID NO:19), DNTVTRS (SEQ ID NO:20), NKISAKD (SEQ ID NO:21), NQDYTKT (SEQ ID NO:22), QADTTKN (SEQ ID NO:23), TNRTSPD (SEQ ID NO:24), SNTTQKT (SEQ ID NO:25), ASDSTKA (SEQ ID NO:26), LANKIQRTDA (SEQ ID NO:27), LANKTTNKDA (SEQ ID NO:28), LATNKIGVTA (SEQ ID NO:29), LAGNLTKGNA (SEQ ID NO:30), LANTVKLSTA (SEQ ID NO:31), LASNTVKAIA (SEQ ID NO:32),

LAASNITKAA (SEQ ID NO:33), LADNTVTRSA (SEQ ID NO:34), LANKISAKDA (SEQ ID NO:35), LANQDYTKTA (SEQ ID NO:36), LATNKIGVTS (SEQ ID NO:37), LATNKIGVTA (SEQ ID NO:38), LAQADTTKNA (SEQ ID NO:39), LATNRTSPDA (SEQ ID NO:40), LASNTTQKTA (SEQ ID NO:41), và LAASDSTKAA (SEQ ID NO:42).

26. AAV biến thể theo mục 23, trong đó trình tự axit amin AAV tương đồng ít nhất 95% với AAV kiểu dại.

27. AAV biến thể theo mục 23, trong đó trình tự axit amin AAV tương đồng ít nhất 99% với AAV kiểu dại.

28. AAV biến thể theo mục 23, trong đó protein capsit là protein capsit thể khảm hoặc protein capsit do tổ tiên truyền lại.

29. Virion virut adeno kết hợp tải tổ hợp (rAAV) lây nhiễm chứa: (a) protein capsit AAV biến thể theo điểm bất kỳ trong số các mục 23-28, và axit nucleic khác loại.

30. rAAV theo điểm 29, trong đó axit nucleic khác loại chứa trình tự nucleotit mã hóa cho tác nhân can thiệp ARN hoặc polypeptit.

31. Phương pháp vận chuyển axit nucleic khác loại đến tế bào đích, bao gồm cho tế bào đích tiếp xúc với virion rAAV theo điểm 29.

32. Phương pháp theo mục 31, trong đó tế bào đích là tế bào cơ xương và/hoặc tế bào cơ tim.

33. Phương pháp theo mục 32, trong đó tế bào tim được chọn từ nhóm bao gồm tế bào cơ tim, cơ tim, nguyên bào sợi tim và tế bào tiền thân tim.

34. Phương pháp theo mục 31, trong đó tế bào đích là *in vitro*.

35. Phương pháp theo mục 31, trong đó tế bào đích là *in vivo*.

36. Axit nucleic được phân lập chứa trình tự nucleotit mã hóa cho protein capsit virut adeno kết hợp (AVV) tương đồng ít nhất 90% với AAV kiểu dại được chọn từ nhóm bao gồm các SEQ ID NO: 1-12 và biến thể AAV; và ii) một hoặc nhiều đột biến thay thế axit amin được chọn từ nhóm bao gồm P363L, P363L+V708I, P363L+E347K, V708I+A593E, V708I+A333S, V708I+S721L, V708I+A593E+N551S, V708I+A593E+K649E, V708I+A593E+S109T, V708I+A593E+S109T+K527Q, A593E+S109T.

37. Tế bào chủ được phân lập chứa axit nucleic theo mục 36.

38. Protein capsit virut adeno kết hợp (AAV) chứa đột biến chèn peptit vào vòng GH trong protein capsit và tùy ý một hoặc nhiều đột biến điểm, trong đó đoạn peptit chèn vào được chọn từ nhóm bao gồm NKIQRTD (SEQ ID NO:13) và LANKIQRTDA (SEQ ID NO:26).

39. Protein capsit AAV biến thể theo điểm 38, chứa đột biến thay thế axit amin V708I.

40. Protein capsit AAV biến thể theo điểm 39, chứa đột biến thay thế axit amin V708I+A593E, V708I+S109T, V708I+T330A, V708I+R588M hoặc V708I+N312K+N449D+N551S+I698V+L735Q.

41. Protein capsit AAV biến thể theo điểm 38, chứa đột biến thay thế axit amin A35P.

42. Protein capsit AAV biến thể theo điểm 38, chứa đột biến thay thế axit amin N312K+N449D+N551S+I698V+L735Q.

43. Protein capsit virut adeno kết hợp (AAV) chứa đột biến chèn peptit vào vòng GH trong protein capsit và tùy ý một hoặc nhiều đột biến điểm, trong đó đoạn peptit chèn vào được chọn từ nhóm bao gồm NKTTNKD (SEQ ID NO:14) và LANKTTNKDA (SEQ ID NO:27).

44. Protein capsit AAV biến thể theo mục 43, chứa đột biến thay thế axit amin V708I.

45. Protein capsit AAV biến thể theo mục 43, chứa đột biến thay thế axit amin V708I+S109T, V708I+W694C, V708I+W606C, hoặc V708I+N312K+N449D+N551S+I698V+L735Q.

46. Protein capsit AAV biến thể theo mục 43, chứa đột biến thay thế axit amin I698V.

47. Protein capsit AAV biến thể theo mục 46, chứa đột biến thay thế axit amin N312K+N449D+N551S+I698V+L735Q.

48. Protein capsit virut adeno kết hợp (AAV) chứa đột biến chèn peptit vào vòng GH trong protein capsit và tùy ý một hoặc nhiều đột biến điểm, trong đó đoạn peptit chèn vào được chọn từ nhóm bao gồm TNKIGVT (SEQ ID NO:15), LATNKIGVTA (SEQ ID NO:28) và LATNKIGVTS (SEQ ID NO:36).

49. Protein capsit AAV biến thể theo mục 48, chứa đột biến thay thế axit amin V708I.

50. Protein capsit AAV biến thể theo mục 49, chứa đột biến thay thế axit amin V708I+N449K, V708I+G222S, hoặc V708I+N312K+N449D+N551S+I698V+L735Q.

51. Protein capsit AAV biến thể theo mục 48, chứa đột biến thay thế axit amin N312K+N449D+N551S+I698V+L735Q.

52. Protein capsit virut adeno kết hợp (AVV) chứa trình tự nêu trong SEQ ID NO:62 hoặc trình tự tương đồng ít nhất 90% với trình tự này, trong đó protein capsit AAV biến thể chứa đột biến thay thế axit amin sau đây liên quan đến capsit AAV5: V229I+A490T+A581T.

53. Protein capsit AAV biến thể theo mục 54, còn chứa thêm đột biến thay thế axit amin Y585S hoặc V447F liên quan đến capsit AAV5.

54. Protein capsit virut adeno kết hợp (AVV) chứa trình tự nêu trong SEQ ID NO:63 hoặc trình tự tương đồng ít nhất 90% với trình tự này, trong đó protein capsit AAV biến thể chứa đột biến thay thế axit amin sau đây liên quan đến capsit AAV5: V229I+A427D+A490T+A581T.

Ví dụ thực hiện sáng chế

Các ví dụ sau đây được đưa ra nhằm cung cấp cho người có hiểu biết trong lĩnh vực kỹ thuật này phần bộc lộ và mô tả hoàn chỉnh để hướng dẫn phương thức điều chế và sử dụng sáng chế, và không nhằm giới hạn phạm vi của những gì các tác giả sáng chế xem là sáng chế của họ cũng không dự định đại diện cho các thử nghiệm dưới đây là tất cả các thử nghiệm hoặc là các thử nghiệm duy nhất được thực hiện. Các nỗ lực đã được thực hiện nhằm đảm bảo độ chính xác đối với các số được sử dụng (ví dụ: số lượng, nhiệt độ, v.v.) nhưng phải tính toán một số lỗi và sai lệch thực nghiệm. Trừ khi có quy định khác, các phần là các phần tính theo trọng lượng, trọng lượng phân tử là trọng lượng phân tử trung bình khối, nhiệt độ là độ C và áp suất ở hoặc gần áp suất khí quyển.

Các phương pháp chung trong sinh hóa phân tử và tế bào có thể được tìm thấy trong các sách giáo khoa tiêu chuẩn như Molecular Cloning: A Laboratory Manual, 3rd Ed. (Sambrook et al., Harbor Laboratory Press 2001); Short Protocols in Molecular Biology, 4th Ed. (Ausubel et al. eds., John Wiley & Sons 1999); Protein Methods (Bollag et al., John Wiley & Sons 1996); Nonviral Vectors for Gene Therapy (Wagner et al. eds., Academic Press 1999); Viral Vectors (Kaplift & Loewy eds., Academic Press 1995); Immunology Methods Manual (I. Lefkovits ed., Academic

Press 1997); và Cell và Tissue Culture: Laboratory Procedures in Biotechnology (Doyle & Griffiths, John Wiley & Sons 1998), phần mô tả của tài liệu này được đưa vào bản mô tả bằng cách viện dẫn. Chất phản ứng, vectơ dòng vô tính, và kit để thao tác di truyền được đề cập trong bản mô tả này đều sẵn có từ các nhà cung cấp thương phẩm như BioRad, Stratagene, Invitrogen, Sigma-Aldrich, và ClonTech.

Ví dụ 1

Tiêm tĩnh mạch và thu hoạch mô. Cá thể khỉ cynomolgus đực (*macaca fascicularis*) 3-10 tuổi và cân nặng ít nhất 3 kg được tiêm theo qua tiêm tĩnh mạch thông qua tĩnh mạch đuôi cho mỗi vòng tuyển chọn. Động vật được gây mê và 1-5ml trong thư viện (trong vòng thứ nhất, thư viện chỉ bao gồm các biến thể được tạo ra sử dụng tất cả các kỹ thuật gây đột biến được mô tả trong Fig.1A; trong mỗi vòng tiếp theo, biến thể được chọn từ vòng trước đó), trong một vài trường hợp được ủ trước với IVIG người trong 30 phút tại 37°C, được cung cấp.

Giết chết nhân đạo bởi nhân viên thú y được đào tạo sử dụng tiêm tĩnh mạch 100 mg natri pentobarbital vào ngày 14±3 hoặc 21±3, tùy thuộc vào việc tuyển chọn. Mô cơ xương và/hoặc cơ tim từ cơ bốn đầu được loại bỏ và ADN được phân lập từ mô. Trong một vài trường hợp, mô tim được chia nhỏ thành một vài vùng: tâm nhĩ, vách liên thất, cơ nhú trái, cơ nhú phải, tâm thất trái, tâm thất phải.

Tiến hóa trực hệ. Quy trình tiến hóa trực hệ được thể hiện trong Fig. 1A-1E. Tóm lại, thư viện capsit virut bao gồm 20+ sự kết hợp độc quyền của kỹ thuật gây đột biến ADN và gen *cap* được tạo ra (Fig. 1A). Virut sau đó được đóng gói (Fig. 1B) — sao cho mỗi hạt bao gồm capsit đột biến bao quanh gen *cap* mã hóa capsit này — và được tinh sạch. Thư viện capsit được đặt dưới áp lực chọn lọc *in vivo*. Mô hoặc vật liệu tế bào quan tâm được thu hoạch để phân lập biến thể AAV mà đã lây nhiễm thành công đích này, và các virut thành công được thu hồi. Các dòng vô tính thành công được làm giàu qua việc chọn lọc lặp lại (Giai đoạn I – Fig. 1D). Gen *cap* đã chọn sau đó trải qua các bước đa dạng hóa lại độc quyền và được làm giàu qua chọn lọc tiếp để làm tăng tiếp sức mạnh của virut (Giai đoạn 2 – Fig. 1D). Các biến thể được nhận diện trong quá trình các giai đoạn Chọn lọc vectơ 1 và 2 chứng minh được khả năng tải nạp tế bào cơ linh trưởng (Fig. 1E).

Thu hồi thành công hệ gen Capsit AAV. Capsit được thu hồi từ mỗi vòng chọn lọc được sử dụng để đóng gói thư viện được tiêm để khơi mào vòng chọn lọc tiếp theo.

Việc thu hồi gen capsit từ mô là sự nội bộ hóa thành công của thư viện vectơ vào mô quan tâm. Việc thu hồi hệ gen virut từ mô cơ xương và cơ tim từ vòng chọn lọc đại diện được thể hiện trong Fig. 2. Các dải trong ô thể hiện sự thu hồi thành công hệ gen virut.

Phân tích trình tự. Trong vòng tuyển chọn 3-4 bao gồm áp lực chọn lọc của vận chuyển trong tĩnh mạch đến mô cơ tim hoặc cơ xương và vòng tuyển chọn 1-2 bao gồm áp lực chọn lọc của vận chuyển trong tĩnh mạch với sự có mặt của kháng thể trung hòa đến mô cơ tim, việc giải trình tự được thực hiện trên dòng tế bào vô tính riêng rẽ trong thư viện để xác định tần suất các biến thể trong quần thể. Các biến thể được đánh giá về sự có mặt của môtip trong dữ liệu giải trình tự. Các biến thể được phân loại thành các môtip dựa trên sự có mặt của sự thay đổi thống nhất (ví dụ, đột biến điểm đặc hiệu hoặc trình tự đoạn peptit chèn vào đặc hiệu ở vị trí thích hợp trong capsit) đã xảy ra trong nhiều trình tự. Môtip là ít nhất 5% quần thể đã được giải trình tự trong hai hoặc nhiều vòng chọn lọc hoặc ít nhất 10% quần thể đã được giải trình tự trong một hoặc nhiều vòng chọn lọc được đại diện trong Fig. 3A (Vòng 4 phân tích trình tự đối với áp lực chọn lọc của việc vận chuyển trong tĩnh mạch đến mô tim), Fig. 3B (Vòng 2 phân tích trình tự đối với áp lực chọn lọc của việc vận chuyển trong tĩnh mạch với sự có mặt của kháng thể trung hòa đến mô tim) và Fig. 3C (Vòng 3 phân tích trình tự đối với áp lực chọn lọc của việc vận chuyển trong tĩnh mạch đến mô cơ xương).

Một số dòng vô tính chọn lọc đã được nhận diện khi tạo ra tính lây nhiễm tăng ở tế bào cơ xương và/hoặc cơ tim được liệt kê trong bảng 1 dưới đây (mỗi dòng vô tính chứa đột biến thay thế và/hoặc đoạn peptit chèn vào đã được nhận diện và nếu không thì tương đồng với trình tự được thể hiện trong SEQ ID NO:2; vòng chọn lọc, số trình tự và tần suất (trong dấu ngoặc đơn) được liệt kê cho mỗi dòng vô tính):

Bảng 1. Biến đổi trình tự axit amin thành protein capsit VP1 AAV khiến cho khả năng lây nhiễm vào tế bào cơ xương và/hoặc cơ tim tăng lên. Đột biến thay thế được liệt kê trong cột 2 dựa trên trình tự axit amin so với AAV2 kiểu dại, tức là không có peptit được chèn vào. "Tim + NAb" trong cột 5 dùng để chỉ rằng biến đổi trình tự axit amin khiến cho khả năng chống lại sự trung hòa bởi kháng thể kháng AAV tăng lên bên cạnh khả năng lây nhiễm vào tế bào cơ tim tăng lên.

Đột biến chèn	Đột biến thay thế	Tim	Tim + NAb	Cơ tim
588~LANKIQRTDA~ (SEQ ID NO:27)	Không có	Vòng 3: 6 (9,68%) Vòng 4: 11 (26,83%)	Vòng 1: 2 (1,41%) Vòng 2: 1 (0,81%)	Vòng 3: 1 (1,23%)
588~LANKIQRTDA~ (SEQ ID NO:27)	+A35P	Vòng 4: 1 (2,44%)	--	--
588~LANKIQRTDA~ (SEQ ID NO:27)	+S109T+V708I	Vòng 3: 1 (1,61%) Vòng 4: 2 (4,88%)	--	--
588~LANKIQRTDA~ (SEQ ID NO:27)	+R588M+V708I	--	Vòng 1: 1 (0,70%)	--
588~LANKIQRTDA~ (SEQ ID NO:27)	+A593E+V708I	Vòng 3: 1 (1,61%)	--	--
588~LANKIQRTDA~ (SEQ ID NO:27)	+V708I	Vòng 3: 13 (20,97%) Vòng 4: 10 (23,26%)	Vòng 1: 1 (0,70%)	--
588~LANKTTNKDA~ (SEQ ID NO:28)	Không có	Vòng 4: 2 (4,88%)	Vòng 1: 10 (7,04%) Vòng 2: 8 (6,50%)	--
588~LANKTTNKDA~ (SEQ ID NO:28)	+S109T+V708I	Vòng 4: 1 (2,44%)	--	--
588~LANKTTNKDA~ (SEQ ID NO:28)	+W694C+V708I	Vòng 4: 1 (2,44%)	--	--
588~LANKTTNKDA~ (SEQ ID NO:28)	+I698V	--	Vòng 1: 1 (0,70%)	--
588~LANKTTNKDA~ (SEQ ID NO:28)	+W606C+V708I	--	Vòng 2: 1 (0,81%)	--

Đột biến chèn	Đột biến thay thế	Tim	Tim + NAb	Cơ tim
588~LANKTTNKDA~ (SEQ ID NO:28)	+V708I	Vòng 3: 6 (9,68%) Vòng 4: 3 (7,32%)	Vòng 1: 4 (2,82%) Vòng 2: 10 (8,13%)	--
588~LATNKIGVTA~ (SEQ ID NO:29)	+V708I	Vòng 4: 1 (2,44%)	--	--
588~LAQADTTKNA~ (SEQ ID NO:39)	Không có	--	Vòng 1: 23 (16,02%) Vòng 2: 21 (17,07%)	--
588~LAQADTTKNA~ (SEQ ID NO:39)	+D213N	--	Vòng 2: 1 (0,81%)	--
588~LAQADTTKNA~ (SEQ ID NO:39)	+G453R	--	Vòng 1: 1 (0,70%)	--
588~LAQADTTKNA~ (SEQ ID NO:39)	+V708I	Vòng 4: 1 (2,44%)	Vòng 1: 3 (2,11%) Vòng 2: 3 (2,44%)	--
588~LAQADTTKNA~ (SEQ ID NO:39)	+P363L	--	Vòng 1: 1 (0,70%)	--
588~LANQDYTKTA~ (SEQ ID NO:36)	Không có	--	Vòng 1: 1 (0,70%)	--
588~LANQDYTKTA~ (SEQ ID NO:36)	+I698V	--	Vòng 2: 2 (1,63%)	--
588~LANQDYTKTA~ (SEQ ID NO:36)	+V708I	--	Vòng 1: 1 (0,70%)	--
588~LATNRTSPDA~ (SEQ ID NO:40)	+V708I	--	Vòng 2: 1 (0,81%)	--
588~LAASDSTKAA~ (SEQ ID NO:42)	Không có	--	--	Vòng 3: 1 (1,23%)

Đột biến chèn	Đột biến thay thế	Tim	Tim + NAb	Cơ tim
588~LAASDSTKAA~ (SEQ ID NO:42)	+V708I	Vòng 3: 2 (3,23%)	--	--
588~LAASNITKAA~ (SEQ ID NO:33)	Không có	--	Vòng 1: 2 (1,41%) Vòng 2: 8 (6,50%)	--
588~LAASNITKAA~ (SEQ ID NO:33)	+V708I	--	Vòng 1: 6 (4,23%) Vòng 2: 11 (8,94%)	--
588~LAGNLTKGNA~ (SEQ ID NO:30)	Không có	Vòng 3: 4 (6,44%)	Vòng 1: 6 (4,23%) Vòng 2: 3 (2,44%)	--
588~LAGNLTKGNA~ (SEQ ID NO:30)	+S109T+V708I	Vòng 3: 2 (3,23%)	--	--
588~LAGNLTKGNA~ (SEQ ID NO:30)	+A139T+P195L	--	Vòng 1: 1 (0,70%)	--
588~LAGNLTKGNA~ (SEQ ID NO:30)	+P363L+V708I	--	Vòng 1: 1 (0,70%)	--
588~LAGNLTKGNA~ (SEQ ID NO:30)	+R651H	--	Vòng 2: 1 (0,81%)	--
588~LAGNLTKGNA~ (SEQ ID NO:30)	+V708I	Vòng 3: 2 (3,23%)	Vòng 1: 1 (0,70%) Vòng 2: 2 (1,63%)	--
588~LAGNLTKGNA~ (SEQ ID NO:30)	+P363L	--	Vòng 1: 1 (0,70%)	--
588~LADNTVTRSA~ (SEQ ID NO:34)	Không có	--	Vòng 1: 9 (6,34%)	--

Đột biến chèn	Đột biến thay thế	Tim	Tim + NAb	Cơ tim
			Vòng 2: 6 (4,88%)	
588~LADNTVTRSA~ (SEQ ID NO:34)	+I698V	--	Vòng 2: 1 (0,81%)	--
588~LADNTVTRSA~ (SEQ ID NO:34)	+V708I	--	Vòng 1: 1 (0,70%) Vòng 2: 2 (1,63%)	--
588~LANTVKLSTA~ (SEQ ID NO:31)	Không có	--	Vòng 1: 3 (2,11%) Vòng 2: 7 (5,69%)	--
588~LANTVKLSTA~ (SEQ ID NO:31)	+V708I	--	Vòng 2: 8 (6,50%)	--
588~LASNTVKAIA~ (SEQ ID NO:32)	Không có	Vòng 3: 2 (3,23%) Vòng 4: 1 (2,44%)	--	--
588~LASNTVKAIA~ (SEQ ID NO:32)	+V708I	Vòng 4: 1 (2,44%)	--	--
588~LATNKIGVTS~ (SEQ ID NO:37)	Không có	Vòng 4: 1 (2,44%)	--	--
588~LASNTTQKTA~ (SEQ ID NO:41)	Không có	--	--	Vòng 3: 2 (2,46%)
588~LANKISAKDA~ (SEQ ID NO:35)	Không có	--	Vòng 2: 3 (2,44%)	--
588~LANKISAKDA~ (SEQ ID NO:35)	+V708I	--	Vòng 2: 2 (1,63%)	--
Không có	P34A	--	--	Vòng 4: 2

Đột biến chèn	Đột biến thay thế	Tim	Tim + NAb	Cơ tim
				(10%)
Không có	P34S	--	Vòng 1: 1 (0,70%)	--
Không có	P64S	--	--	Vòng 4: 1 (5,00%)
Không có	S109T+P235S	--	--	Vòng 4: 1 (5,00%)
Không có	Q120R	--	--	Vòng 4: 1 (5,00%)
Không có	A193V	--	--	Vòng 3: 1 (1,23%)
Không có	T277N	--	--	Vòng 4: 1 (5,00%)
Không có	P351L	--	--	Vòng 4: 1 (5,00%)
Không có	P363L	--	Vòng 1: 13 (9,15%)	Vòng 3: 34 (41,98%) Vòng 4: 4 (20,00%)
Không có	P363L+E347K	--	--	Vòng 3: 1 (1,23%)
Không có	P363L+V708I	--	Vòng 1: 2 (1,41%)	--
Không có	S427T+I698V	--	Vòng 1: 1 (0,70%)	--
Không có	Q440K	--	--	Vòng 3: 1 (1,23%)
Không có	Y444F	--	--	Vòng 4: 1 (5,00%)
Không có	N449D	--	--	Vòng 4: 1

Đột biến chèn	Đột biến thay thế	Tim	Tim + NAb	Cơ tim
				(5,00%)
Không có	T568N	--	--	Vòng 3: 1 (1,23%)
Không có	A593E	Vòng 4: 3 (7,32%)	Vòng 1: 1 (0,70%)	Vòng 3: 31 (38,27%) Vòng 4: 2 (10%)
Không có	S109T+A593E	--	--	Vòng 3: 2 (2,47%)
Không có	S109T+K527Q+ A593E+V708I	Vòng 3: 1 (1,61%)	--	--
Không có	S109T+A593E+ V708I	Vòng 3: 1 (1,61%)	--	--
Không có	A593E+N551S+ V708I	--	--	Vòng 3: 1 (1,23%)
Không có	A593E+K649E+ V708I	Vòng 3: 1 (1,61%)	--	--
Không có	A593E+V708I	Vòng 3: 12 (19,35%)	Vòng 1: 5 (3,52%)	Vòng 3: 1 (1,23%)
Không có	I698V	--	--	Vòng 4: 1 (5,00%)
Không có	V708I	Vòng 3: 6 (9,68%) Vòng 4: 2 (4,88%)	Vòng 1: 10 (7,04%)	Vòng 3: 1 (1,23%) Vòng 4: 1 (5,00%)
Không có	V708I+A333S	--	--	Vòng 3: 1 (1,23%)
Không có	V708I+S721L	Vòng 3: 1 (1,61%)	--	--
Không có	V708I+L735V	--	--	Vòng 3: 1

Đột biến chèn	Đột biến thay thế	Tim	Tim + NAb	Cơ tim
				(1,23%)

Cũng được nhận diện dưới dạng capsit có tính lây nhiễm tăng của một hoặc nhiều tế bào cơ tim và tăng khả năng chống lại sự trung hòa bởi kháng thể kháng AAV là các thể khảm sau đây:

Thể khảm với (i) axit amin 1-129 trong AAV6 và (ii) axit amin 130-725 trong AAV5 và có đột biến thay thế axit amin sau đây liên quan đến AAV5: V229I+A490T+A581T (trình tự nêu trong SEQ ID NO:62).

Thể khảm với (i) axit amin 1-61 trong AAV2, (ii) axit amin 62-129 trong AAV6, và (iii) axit amin 130-725 trong AAV5 và có đột biến thay thế axit amin sau đây liên quan đến AAV5: V229I+A490T+A581T (trình tự nêu trong SEQ ID NO:63).

Thể khảm với (i) axit amin 1-129 trong AAV6 và (ii) axit amin 130-725 trong AAV5 và có đột biến thay thế axit amin sau đây liên quan đến AAV5: V229I+A490T+A581T+Y585S.

Thể khảm bao gồm (i) axit amin 1-129 trong AAV6 và (ii) axit amin 130-725 trong AAV5 và có đột biến thay thế axit amin sau đây liên quan đến AAV5: V229I+A447F+A490T+A581T.

Virion biến thể AAV được mô tả trong bản mô tả này có thể kết hợp các tham số thiết kế hợp lý, các dấu hiệu, biến đổi, ưu điểm và thay đổi đã biết đối với những người có hiểu biết trung bình trong lĩnh vực thiết kế vectơ virut AAV.

Ví dụ 2

Tính hướng tế bào của virion AAV tái tổ hợp chứa biến thể AAV mới LANKIQRTDA+V708I (SEQ ID NO:43), LANKTTNKDA+V708I (SEQ ID NO:48), và LATNKIGVTA+V708I (SEQ ID NO:46) đến tế bào cơ tim được đánh giá *in vitro* sử dụng tế bào cơ tim được tạo ra từ tế bào gốc phôi người (ESC).

Virion AAV tái tổ hợp chứa capsit AAV1, capsit AAV2, capsit AAV9, capsit biến thể mới LANKIQRTDA+V708I, capsit biến thể mới LANKTTNKDA+V708I hoặc capsit biến thể mới LATNKIGVTA+V708I và hệ gen chứa gen chuyển protein huỳnh quang xanh (EGFP) được liên kết có điều khiển với vùng gen khởi đầu CAG (lần lượt là AAV1.CAG.EGFP, AAV2.CAG.EGFP,AAV9.CAG.EGFP, LANKIQRTDA+V708I.CAG.EGFP, LANKTTNKDA+V708I.CAG.EGFP, và

LATNKIGVTA+V708I.CAG.GFP) được sản xuất sử dụng phương pháp tiêu chuẩn. Tế bào cơ tim được tạo ra từ dòng tế bào gốc phôi người, ESI-017, bằng cách thay đổi tín hiệu Wnt sử dụng phân tử nhỏ. Sau 14 ngày cảm ứng trung bì tim, môi trường nuôi cấy tiếp tục được làm giàu cho các tế bào cơ tim do thiếu glucozơ. Sau gần 24 ngày biệt hóa, phần lớn các tế bào biểu hiện chỉ thị tế bào cơ tim, Troponin T tim (cTnT) và chỉ thị đặc hiệu tâm thất, MLC-2V. Các tế bào cơ tim được tạo ra được đánh giá để biểu hiện protein liên kết khe Connexin 43, biến thiên điện thế màng, xử lý canxi, va chức năng co bóp để đảm bảo rằng tế bào cơ tim được tạo ra đạt đến trạng thái trưởng thành trước khi biểu thị đặc điểm của vectơ.

Liên quan đến AAV1, AAV2, và AAV9, các biến thể LANKIQRTDA+V708I, LANKTTNKDA+V708I, và LATNKIGVTA+V708I được tạo ra để tăng đáng kể hiệu quả tải nạp và sự biểu hiện của gen chuyển trong môi trường nuôi cấy tế bào tim sáu ngày sau khi lây nhiễm khi được xác định bằng miễn dịch huỳnh quang (Fig.6A), đếm dòng chảy tế bào (Fig. 6B) và phân tích Western blot (Fig. 6C-D). Ngoài ra, liên quan đến AAV1, AAV2, và AAV9, LANKIQRTDA+V708I, LANKTTNKDA+V708I, và LATNKIGVTA+V708I được tạo ra để bắt đầu nhanh hơn sự biểu hiện gen trong môi trường nuôi cấy tế bào cơ tim người, như được xác định bằng miễn dịch huỳnh quang (Fig. 6E). Liên quan đến AAV8 và AAV9, cái mà biểu hiện tính hướng tế bào cơ xương và cơ tim, số lượng đơn vị lây nhiễm trên hệ gen virut được cung cấp cao hơn gấp nhiều lần so với LANKIQRTDA+V708I và LANKTTNKDA+V708I (Fig 10A). Nghiên cứu này chứng minh khả năng vượt trội của NKIQRTD (SEQ ID NO:13)-, NKTTNKD (SEQ ID NO:14)-, và TNKIGVT (SEQ ID NO:15)-chứa biến thể capsit AAV trong việc vận chuyển các gen đến tế bào tim.

Ví dụ 3

Tính hướng tế bào của virion AAV tái tổ hợp chứa biến thể AAV mới là thể khảm AAV6/AAV5 đến tế bào cơ tim được đánh giá *in vitro* sử dụng tế bào cơ tim được tạo ra từ tế bào gốc phôi người (ESC).

Virion AAV tái tổ hợp chứa capsit AAV1, capsit AAV8 hoặc capsit AAV9, hoặc capsit biến thể thể khảm AAV6/AAV5 mới (SEQ ID NO:62) và hệ gen chứa gen chuyển protein phát huỳnh quang xanh (EGFP) được liên kết có điều khiển với vùng gen khởi đầu CAG (lần lượt là AAV1.CAG.EGFP, AAV8.CAG.EGFP, AAV9.CAG.EGFP, thể khảm AAV6/AAV5.CAG.EGFP) được tạo ra sử dụng phương

pháp tiêu chuẩn. Các tế bào cơ tim được tạo ra từ dòng tế bào gốc phôi người, ESI-017, bằng cách thay đổi tín hiệu Wnt sử dụng các phân tử nhỏ. Sau 14 ngày cảm ứng trung bì tim, môi trường nuôi cấy tiếp tục được làm giàu cho các tế bào cơ tim do thiếu glucozơ. Sau gần 24 ngày biệt hóa, phần lớn các tế bào biểu hiện chỉ thị tế bào cơ tim, Troponin T tim (cTnT) và chỉ thị đặc hiệu tâm thất, MLC-2V. Các tế bào cơ tim được tạo ra được đánh giá để biểu hiện protein liên kết khe Connexin 43, biến thiên điện thế màng, xử lý canxi, và chức năng co bóp để đảm bảo rằng tế bào cơ tim được tạo ra đạt đến trạng thái trưởng thành trước khi biểu thị đặc điểm của vectơ.

Liên quan đến AAV1, AAV8, và AAV9, thể khảm AAV6/AAV5 được tạo ra để tăng đáng kể hiệu quả tải nạp và sự biểu hiện của gen chuyển trong môi trường nuôi cấy tế bào tim sáu ngày sau khi lây nhiễm khi được xác định bằng miễn dịch huỳnh quang (Fig.76A), đếm dòng chảy tế bào (Fig. 7B) và phân tích Western blot (Fig. 7C-D). Ngoài ra, liên quan đến AAV8, thể khảm AAV6/AAV5 được tạo ra để bắt đầu nhanh hơn sự biểu hiện gen trong môi trường nuôi cấy tế bào cơ tim người, như được xác định bằng miễn dịch huỳnh quang (Fig. 7E). Liên quan đến AAV8 và AAV9, cái mà biểu hiện tính hướng tế bào cơ xương và cơ tim, số lượng đơn vị lây nhiễm trên hệ gen virut được cung cấp cao hơn gấp nhiều lần so với thể khảm AAV6/AAV5 (Fig 10A). Nghiên cứu này chứng minh khả năng vượt trội của SEQ ID NO:62-chứa biến thể capsit AAV để vận chuyển các gen đến tế bào tim.

Ví dụ 4

Tính hướng tế bào của virion AAV tái tổ hợp chứa biến thể AAV LANKIQRTDA+V708I,LANKTTNKDA+V708I, và thể khảm AAV6/AAV5 đến tế sợi tơ cơ xương được đánh giá *in vitro* sử dụng sợi tơ cơ xương được tạo ra từ nguyên bào cơ người sơ cấp.

Virion AAV tái tổ hợp chứa capsit AAV8 hoặc capsit AAV9, hoặc capsit biến thể mới LANKIQRTDA+V708I, capsit biến thể mới LANKTTNKDA+V708I, thể khảm AAV6/AAV5 và hệ gen chứa gen chuyển protein phát huỳnh quang xanh (EGFP) được liên kết có điều khiển với vùng gen khởi đầu CAG (lần lượt là AAV8.CAG.EGFP, AAV9.CAG.EGFP, LANKIQRTDA+V708I.CAG.EGFP, LANKTTNKDA+V708I.CAG.EGFP, và thể khảm AAV6/AAV5.CAG.GFP) được tạo ra sử dụng phương pháp tiêu chuẩn. Các sợi tơ cơ xương được tạo ra từ nguyên bào cơ xương người sơ cấp thu được từ nam giới 51 tuổi khỏe mạnh (Cook Myosites).

Nguyên bào sợi cơ được biệt hóa 30 ngày để tạo ra sợi cơ xương đa nhân trưởng thành. Sợi tơ cơ xương được tạo ra được đánh giá biểu hiện của chuỗi nặng myosin và dystrophin để đảm bảo rằng phần lớn sợi tơ cơ xương đạt đến trạng thái trưởng thành trước khi biểu thị đặc điểm của vectơ.

Liên quan đến AAV8, và AAV9, LANKIQRTDA+V708I, LANKTTNKDA+V708I và thể khảm AAV6/AAV5 được tạo ra để tăng đáng kể hiệu quả tải nạp và sự biểu hiện của gen chuyển trong môi trường nuôi cấy tế bào tim sáu ngày sau khi lây nhiễm khi được xác định bằng miễn dịch huỳnh quang (Fig.8A), đếm dòng chảy tế bào (Fig. 8B). Ngoài ra, liên quan đến AAV8, và AAV9, LANKIQRTDA+V708I và LANKTTNKDA+V708I được tạo ra để bắt đầu nhanh hơn sự biểu hiện gen trong môi trường nuôi cấy tế bào cơ tim người, như được xác định bằng miễn dịch huỳnh quang (Fig. 8C). Liên quan đến AAV8 và AAV9, cái mà biểu hiện tính hướng tế bào cơ xương và cơ tim, số lượng đơn vị lây nhiễm trên hệ gen virut được cung cấp cao hơn gấp nhiều lần so với LANKIQRTDA+V708I, LANKTTNKDA+V708I, thể khảm AAV6/AAV5 (Fig 10B). Nghiên cứu này chứng minh khả năng vượt trội của NKIQRTD (SEQ ID NO:13)-, NKTTNKD (SEQ ID NO:14)-, và SEQ ID NO:62-chứa biến thể capsit AAV để vận chuyển các gen đến sợi tơ cơ xương.

Ví dụ 5

Tính hướng tế bào của virion AAV tái tổ hợp chứa biến thể AAV mới LANKIQRTDA+V708I, LANKTTNKDA+V708I, và thể khảm AAV6/AAV5 đến tế bào cơ xương nguyên sơ được đánh giá *in vitro* sử dụng tế bào cơ xương nguyên bản được tạo ra từ tế bào gốc vạn năng cảm ứng người có nguồn gốc từ nguyên bào sợi (FB-iPSC) hoặc tế bào gốc phôi người (ESC).

Virion AAV tái tổ hợp chứa capsit AAV9, capsit biến thể mới LANKIQRTDA+V708I, capsit biến thể mới LANKTTNKDA+V708I, hoặc capsit biến thể thể khảm AAV6/AAV5 mới (SEQ ID NO:62) và hệ gen chứa gen chuyển protein phát huỳnh quang xanh (EGFP) được liên kết có điều khiển với vùng gen khởi đầu CAG (lần lượt là AAV8.CAG.EGFP, AAV9.CAG.EGFP, LANKIQRTDA+V708I.CAG.EGFP, LANKTTNKDA+V708I.CAG.EGFP, và thể khảm AAV6/AAV5.CAG.GFP được tạo ra sử dụng phương pháp tiêu chuẩn. Các tế bào cơ xương nguyên sơ được tạo ra từ dòng tế bào gốc phôi người, ESI-017 (ESI-

BIO) sau khi biệt hóa được mô tả trong Shelton et al. Methods, 2016 cùng với biến đổi nhỏ. Sau gần 40 ngày biệt hóa, cắt giới hạn tuyến tính tế bào cơ xương nguyên sơ được xác nhận bằng biểu hiện của PAX7, và MyoD trong phần lớn tế bào trước khi sử dụng môi trường nuôi cấy để biểu thị đặc điểm của vectơ.

Liên quan đến AAV9, LANKIQRTDA+V708I, LANKTTNKDA+V708I, và thể khảm AAV6/AAV5 được tạo ra để tăng đáng kể hiệu quả tải nạp và sự biểu hiện của gen chuyển trong môi trường nuôi cấy tế bào tim sáu ngày sau khi lây nhiễm khi được xác định bằng miễn dịch huỳnh quang (Fig.9A), đếm dòng chảy tế bào (Fig. 9B). Nghiên cứu này chứng minh khả năng vượt trội của NKIQRTD (SEQ ID NO:13)-, NKTTNKD (SEQ ID NO:14)-, và SEQ ID NO:62-chứa biến thể capsit AAV để vận chuyển các gen đến tế bào tim.

Ví dụ 6

Tiến hóa trực hệ được dùng để phát hiện ra các biến thể virut adeno kết hợp (AAV) mới có khả năng vận chuyển gen vượt trội đến tế bào cơ xương và cơ tim sau khi tiêm trong tĩnh mạch, đường dùng có ưu điểm đáng kể so với các phương pháp vận chuyển gen khác đến cơ xương và cơ tim người (Ví dụ 1). Tính hướng tế bào sau khi tiêm trong cơ virion AAV tái tổ hợp chứa biến thể AAV mới bao gồm đột biến thay thế V708I và peptit LANKIQRTDA(SEQ ID NO:27) được thêm vào giữa axit amin 587 và 588 (LANKIQRTDA+V708I; SEQ ID NO:43) được đánh giá *in vivo* ở chuột là ví dụ đại diện về khả năng của virion rAAV chứa NKIQRTD(SEQ ID NO:13)-chứa biến thể capsit AAV trong việc tải nạp tế bào cơ.

Virion AAV tái tổ hợp bao gồm capsit biến thể mới LANKIQRTDA+V708I và hệ gen bao gồm gen chuyển luciferaza được liên kết có điều khiển với vùng gen khởi đầu CAG (LANKIQRTDA+V708I.CAG.luciferaza) được tạo ra bằng cách sử dụng phương pháp tiêu chuẩn. Chuột B6 Albino (C57BL/6) được tiêm qua đường tĩnh mạch đuôi với $2x10^{12}$ vg, và sự tải nạp được đánh giá lúc sống hình ảnh luciferaza và sau khi chết bằng hoạt động của luciferaza ở mô. Hình ảnh của luciferaza tại ngày 14 (bên trái) và ngày 28 (bên phải) sau khi tiêm chứng minh rằng capsit biến thể AAV mới LANKIQRTDA+V708I có thể tải nạp tế bào chuột *in vivo* (Fig 11A). Hoạt động của luciferaza trong tim, cơ hoành và cơ bốn đầu vào 56 ngày sau cung cấp chứng minh rằng capsit biến thể AAV mới LANKIQRTDA+V708I có thể tải nạp cơ xương và cơ tim *in vivo* (Fig 11B).

Nghiên cứu này chứng minh rằng việc vận chuyển gen bằng biến thể chứa NKIQRTD (SEQ ID NO:13) sau khi dùng một trong số một vài đường dùng có thể chấp nhận lâm sàng. Hiệu quả tương tự có thể đạt được với các biến thể khác chứa motip chèn thêm peptit này. Tương tự như vậy, hiệu quả tương tự có thể thu được với các biến thể khác được mô tả ở đây đã được xác định bằng cách sử dụng phương pháp tiến hóa tương tự.

Ví dụ 7.

Tiến hóa trực hệ được dùng để phát hiện ra các biến thể virut adeno kết hợp (AAV) mới có khả năng vận chuyển gen vượt trội đến tế bào cơ xương và cơ tim sau khi tiêm trong tĩnh mạch, đường dùng có ưu điểm đáng kể so với các phương pháp vận chuyển gen khác đến cơ xương và cơ tim người (Ví dụ 1). Tính hướng tế bào sau khi tiêm trong cơ virion AAV tái tổ hợp chứa biến thể AAV mới bao gồm đột biến thay thế V708I và peptit LANKIQRTDA(SEQ ID NO:27) được chèn vào giữa axit amin 587 và 588 (LANKIQRTDA+V708I; SEQ ID NO:43) được đánh giá *in vivo* ở linh trưởng không phải người (NHP) là ví dụ đại diện về khả năng của biến thể rAAV chứa biến thể capsit AAV chứa NKIQRTD(SEQ ID NO:13) trong việc tải nạp tế bào cơ.

Virion AAV tái tổ hợp bao gồm capsit biến thể mới LANKIQRTDA+V708I và hệ gen bao gồm gen chuyển protein phát huỳnh quang xanh (GFT) được liên kết có điều khiển với vùng gen khởi đầu CAG (LANKIQRTDA+V708I.CAG.GFP) được tạo ra bằng cách sử dụng phương pháp tiêu chuẩn. Khỉ đuôi dài Cynomolgus được tiêm trong cơ với ba liều vectơ vào trong các vị trí tại vùng bụng sau với $1x10^{11}$ vg và tải nạp tế bào cơ xương được đánh giá sau khi chết bằng hình ảnh miễn dịch huỳnh quang. Các hình ảnh đại diện của haemotoxylin và eosin (H&E) và nhuộm màu kháng thể kháng GFP của mặt cắt ngang vị trí sinh thiết gần nhất ở độ phóng đại 2x, 4x, và 20x chứng minh rằng capsit biến thể AAV mới LANKIQRTDA+V708I có thể tải nạp tế bào cơ xương linh trưởng *in vivo* (Fig. 12A). Hình ảnh đại diện của và eosin (H&E) và nhuộm kháng thể kháng GFP của mặt cắt dọc vị trí sinh thiết gần nhất ở độ phóng đại 2x, 4x, và 20x chứng minh rằng capsit biến thể AAV mới LANKIQRTDA+V708I có thể tải nạp tế bào cơ xương linh trưởng *in vivo* (Fig. 12B).

Nghiên cứu này chứng minh rằng việc vận chuyển gen bằng biến thể chứa NKIQRTD (SEQ ID NO:13) sau khi dùng một trong số một vài đường dùng có thể chấp nhận lâm sàng. Hiệu quả tương tự có thể đạt được với các biến thể khác chứa

motip chèn thêm peptit này. Tương tự như vậy, hiệu quả tương tự có thể thu được với các biến thể khác được mô tả ở đây đã được xác định bằng cách sử dụng phương pháp tiến hóa tương tự.

Các phần mô tả trên đây chỉ nhằm minh họa nguyên lý theo sáng chế. Cần phải hiểu rằng người có hiểu biết trong lĩnh vực kỹ thuật này sẽ có thể sẽ có thể đưa ra các sự sắp xếp khác nhau, mặc dù không được mô tả hoặc thể hiện rõ ràng trong bản mô tả này, thể hiện nguyên lý theo sáng chế và nằm trong nội dung và phạm vi của sáng chế. Hơn nữa, tất cả các ví dụ và ngôn ngữ có điều kiện được thể hiện trong bản mô tả này chủ yếu nhằm hỗ trợ người đọc hiểu được các nguyên lý theo sáng chế và các khái niệm do các tác giả sáng chế đóng góp để phát triển lĩnh vực kỹ thuật này và được hiểu là không giới hạn ở các ví dụ và điều kiện cụ thể như vậy.

Ngoài ra, tất cả các phần đề cập trong bản mô tả này nói đến nguyên lý, khía cạnh, và các phương án theo sáng chế cũng như các ví dụ cụ thể của chúng, được dự định bao hàm cả các dạng tương đương về cấu trúc và chức năng của chúng. Ngoài ra, cũng dự định rằng các dạng tương đương như vậy bao hàm cả các dạng tương đương hiện đã biết và các dạng tương đương phát triển trong tương lai, tức là, yếu tố bất kỳ đã phát triển để thực hiện cùng chức năng, bất kể cấu trúc gì. Vì vậy, phạm vi của sáng chế không dự định giới hạn ở các phương án được lấy làm ví dụ được thể hiện và đề cập trong bản mô tả này. Đúng hơn, phạm vi và nội dung của sáng chế được thể hiện bởi các điểm yêu cầu bảo hộ kèm theo.

YÊU CẦU BẢO HỘ

1. Protein capsit virut adeno kết hợp (AAV) biến thể chứa đột biến chèn peptit vào vòng GH của protein capsit so với protein capsit AAV mẹ tương ứng, trong đó protein capsit khiến cho khả năng lây nhiễm của virion AAV tái tổ hợp (rAAV) vào tế bào cơ tăng lên so với khả năng lây nhiễm vào tế bào cơ bởi virion AAV chứa protein capsit AAV mẹ tương ứng, trong đó peptit chèn vào bao gồm trình tự axit amin LANKIQRTDA (SEQ ID NO:27) hoặc bao gồm trình tự axit amin LANKTTNKDA (SEQ ID NO:28), và trong đó vị trí chèn peptit vào nằm giữa axit amin 587 và 588 trong VP1 của AAV2 (SEQ ID NO:2).

2. Protein capsit AAV biến thể theo điểm 1, trong đó protein capsit AAV chứa một hoặc nhiều đột biến thay thế axit amin so với VP1 của AAV2 (SEQ ID NO:2).

3. Protein capsit AAV biến thể theo điểm 2, trong đó peptit chèn vào bao gồm trình tự axit amin LANKTTNKDA (SEQ ID NO:28) và trong đó protein capsit AAV còn bao gồm thêm một hoặc nhiều đột biến thay thế axit amin V708I, đột biến thay thế axit amin S109T, đột biến thay thế axit amin W694C và đột biến thay thế axit amin W606C so với VP1 của AAV2 (SEQ ID NO:2).

4. Protein capsit AAV biến thể theo điểm 2, trong đó protein capsit AAV gần như chỉ bao gồm trình tự axit amin nêu trong SEQ ID NO:48.

5. Protein capsit AAV biến thể theo điểm 2, trong đó peptit chèn vào bao gồm trình tự axit amin LANKIQRTDA (SEQ ID NO:27) và trong đó protein capsit AAV còn bao gồm thêm một hoặc nhiều đột biến thay thế axit amin V708I, đột biến thay thế axit amin S109T, đột biến thay thế axit amin R588M và đột biến thay thế axit amin A593E so với VP1 của AAV2 (SEQ ID NO:2).

6. Protein capsit AAV biến thể theo điểm 2, trong đó protein capsit AAV gần như chỉ bao gồm trình tự axit amin nêu trong SEQ ID NO:43.

7. Protein capsit AAV biến thể theo điểm bất kỳ trong số các điểm từ 1 đến 6, trong đó protein capsit còn khiến cho khả năng chống lại sự trung hòa bởi kháng thể trung hòa của virion rAAV lây nhiễm tăng lên so với AAV chứa protein capsit AAV mẹ tương ứng.

8. Axit nucleic được phân lập chứa trình tự nucleotit mà mã hóa cho protein capsit

AAV biến thể theo điểm bất kỳ trong số các điểm từ 1 đến 7.

9. Virion AAV tái tổ hợp (rAAV) lây nhiễm chứa protein capsit AAV biến thể theo điểm bất kỳ trong số các điểm từ 1 đến 7.

10. Virion rAAV theo điểm 9, còn chứa thêm axit nucleic khác loại chứa trình tự nucleotit mã hóa cho sản phẩm gen.

11. Virion rAAV theo điểm 10, trong đó sản phẩm gen là protein, ARN can thiệp nhỏ, ARN đối mã, microARN, ARN kẹp tóc ngắn hoặc ARN can thiệp nhỏ.

12. Virion rAAV theo điểm 11, trong đó virion rAAV chứa axit nucleic khác loại mã hóa cho sản phẩm gen được chọn từ alpha galactosidaza A (GLA), frataxin (FXN), dystrophin (DMD) hoặc mảnh chức năng của chúng, axit alpha glucosidaza (GAA), và glycogen phosphorylaza, cơ (PYGM).

13. Virion rAAV theo điểm 12, trong đó protein capsit AAV biến thể bao gồm trình tự axit amin nêu trong SEQ ID NO:48 và trong đó trình tự nucleotit mã hóa GLA được liên kết có điều khiển với vùng gen khởi đầu CAG.

14. Virion rAAV theo điểm 12, trong đó protein capsit AAV biến thể bao gồm trình tự axit amin nêu trong SEQ ID NO:48 và trong đó trình tự nucleotit mã hóa glycogen phosphorylaza, cơ (PYGM) được liên kết có điều khiển với vùng gen khởi đầu CAG.

15. Virion rAAV theo điểm 12, trong đó protein capsit AAV biến thể bao gồm trình tự axit amin nêu trong SEQ ID NO:48 và trong đó trình tự nucleotit mã hóa a frataxin protein được liên kết có điều khiển với vùng gen khởi đầu CAG.

16. Virion rAAV theo điểm 12, trong đó protein capsit AAV biến thể bao gồm trình tự axit amin nêu trong SEQ ID NO:48 và trong đó trình tự nucleotit mã hóa protein GAA được liên kết có điều khiển với vùng gen khởi đầu CBA.

17. Dược phẩm chứa virion rAAV theo điểm bất kỳ trong số các điểm từ 12 đến 16 và tá dược dược dụng.

Fig. 1

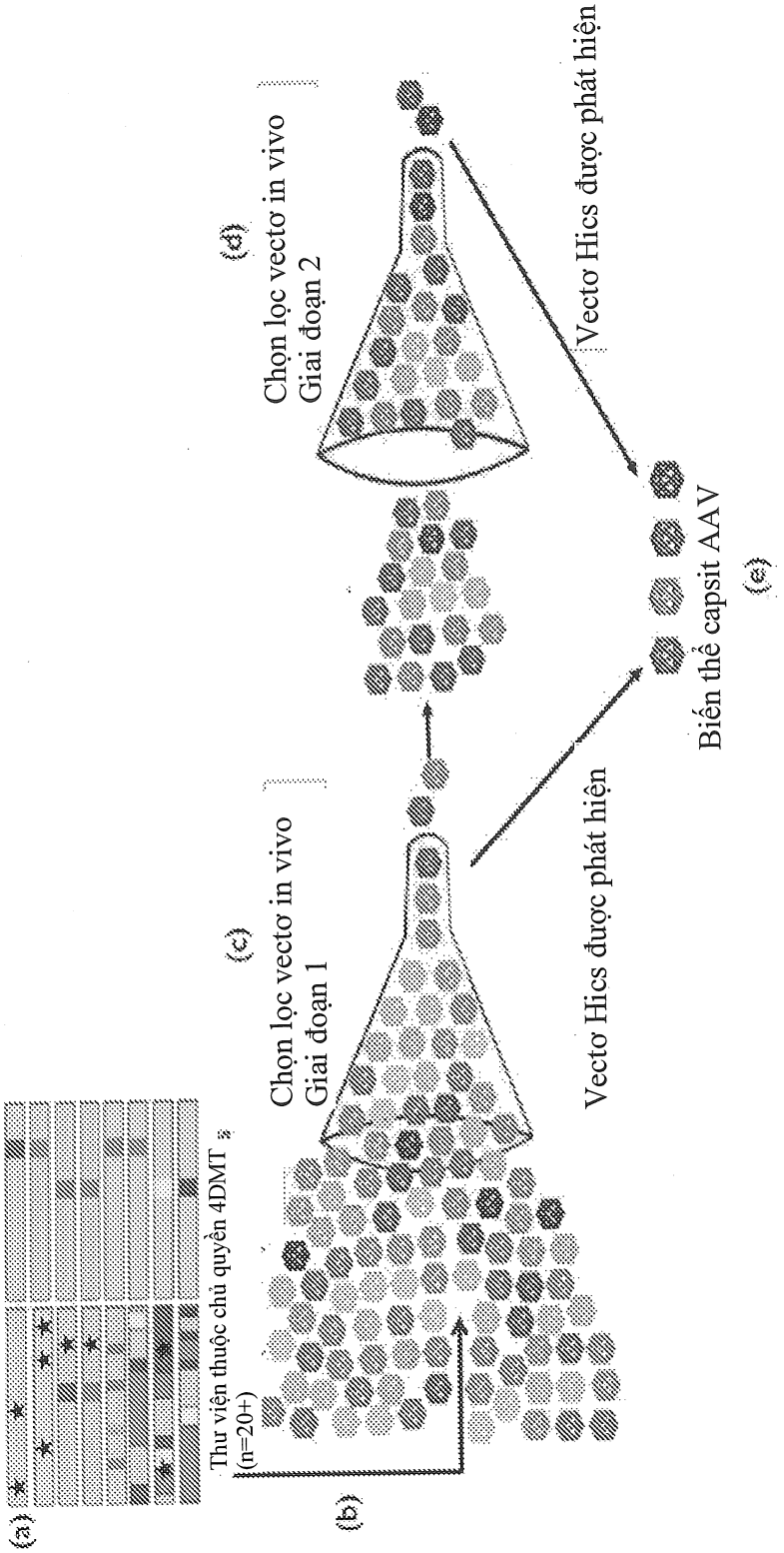

(a)
Thư viện thuộc chủ quyền 4DMT
(n=20+)
(b)
(c)
Chọn lọc vectơ in vivo
Giai đoạn 1
Vectơ Hics được phát hiện
(d)
Chọn lọc vectơ in vivo
Giai đoạn 2
Vectơ Hics được phát hiện
Biến thể capsit AAV
(e)

Fig. 2

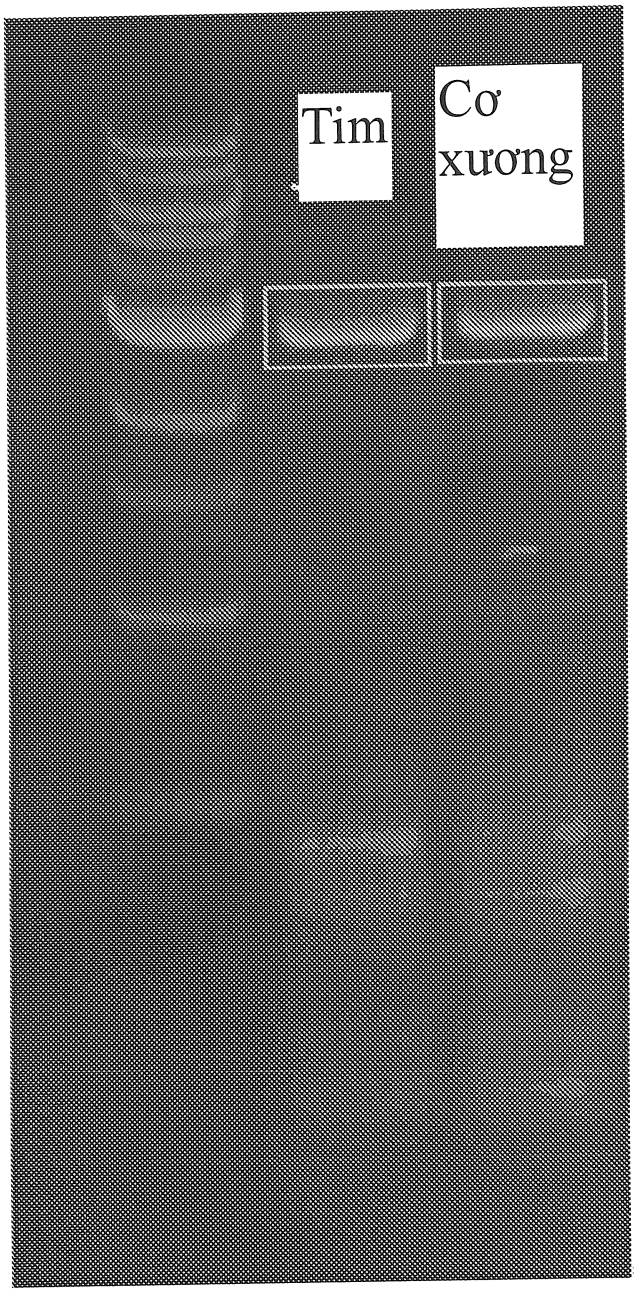


Fig 3A

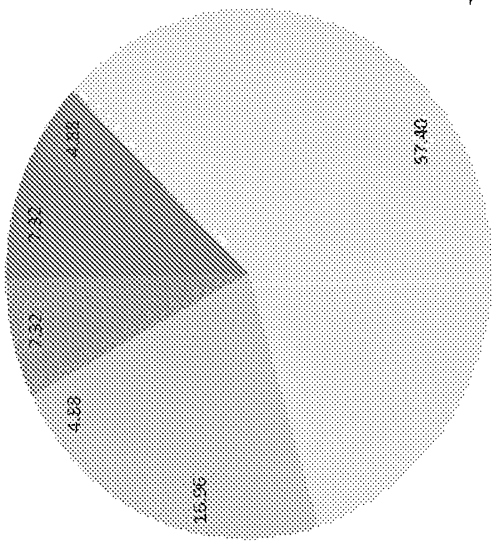


Fig 3B

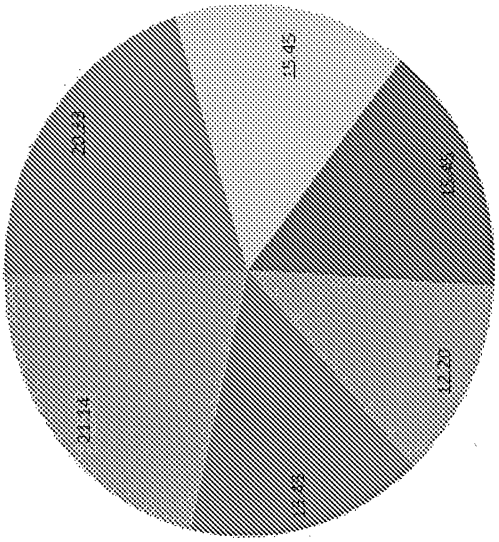
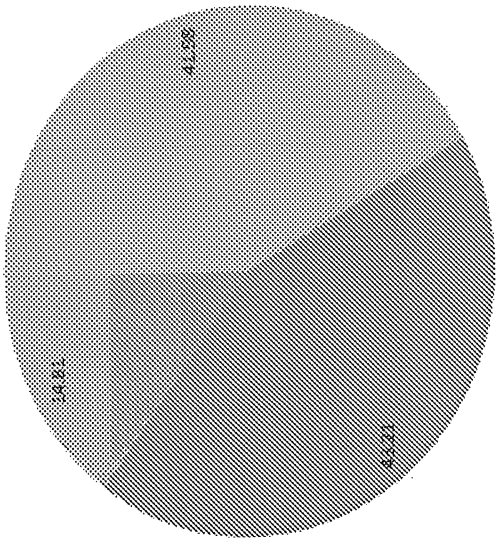


Fig 3C



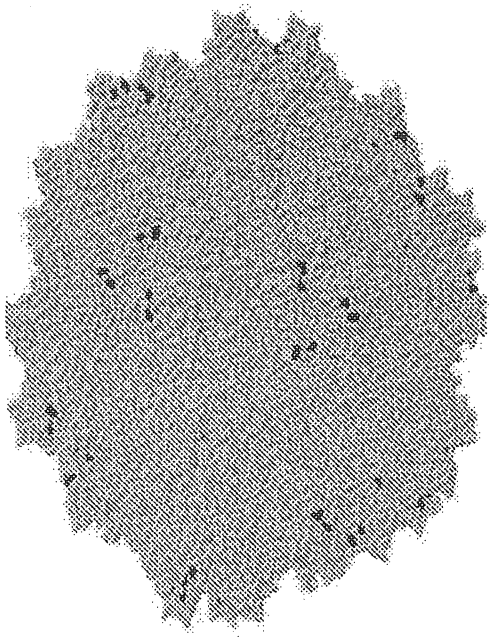


Fig. 4B

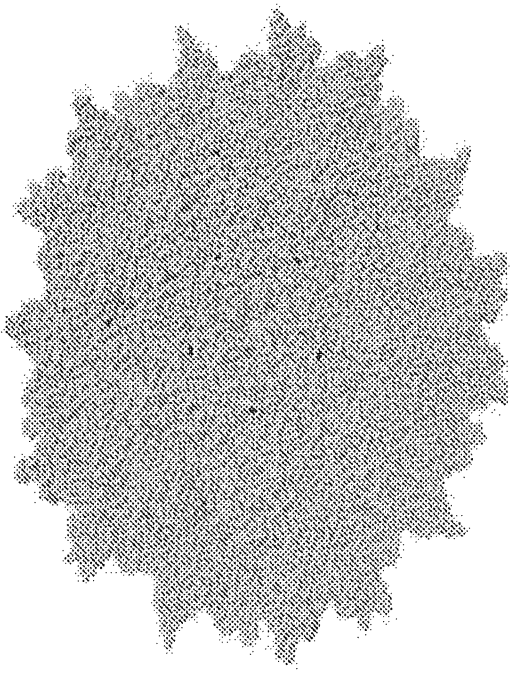


Fig. 4C

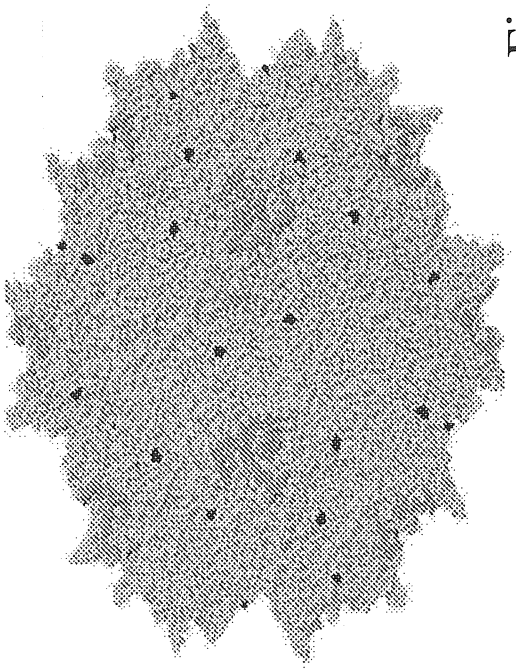


Fig. 4A

Fig. 5

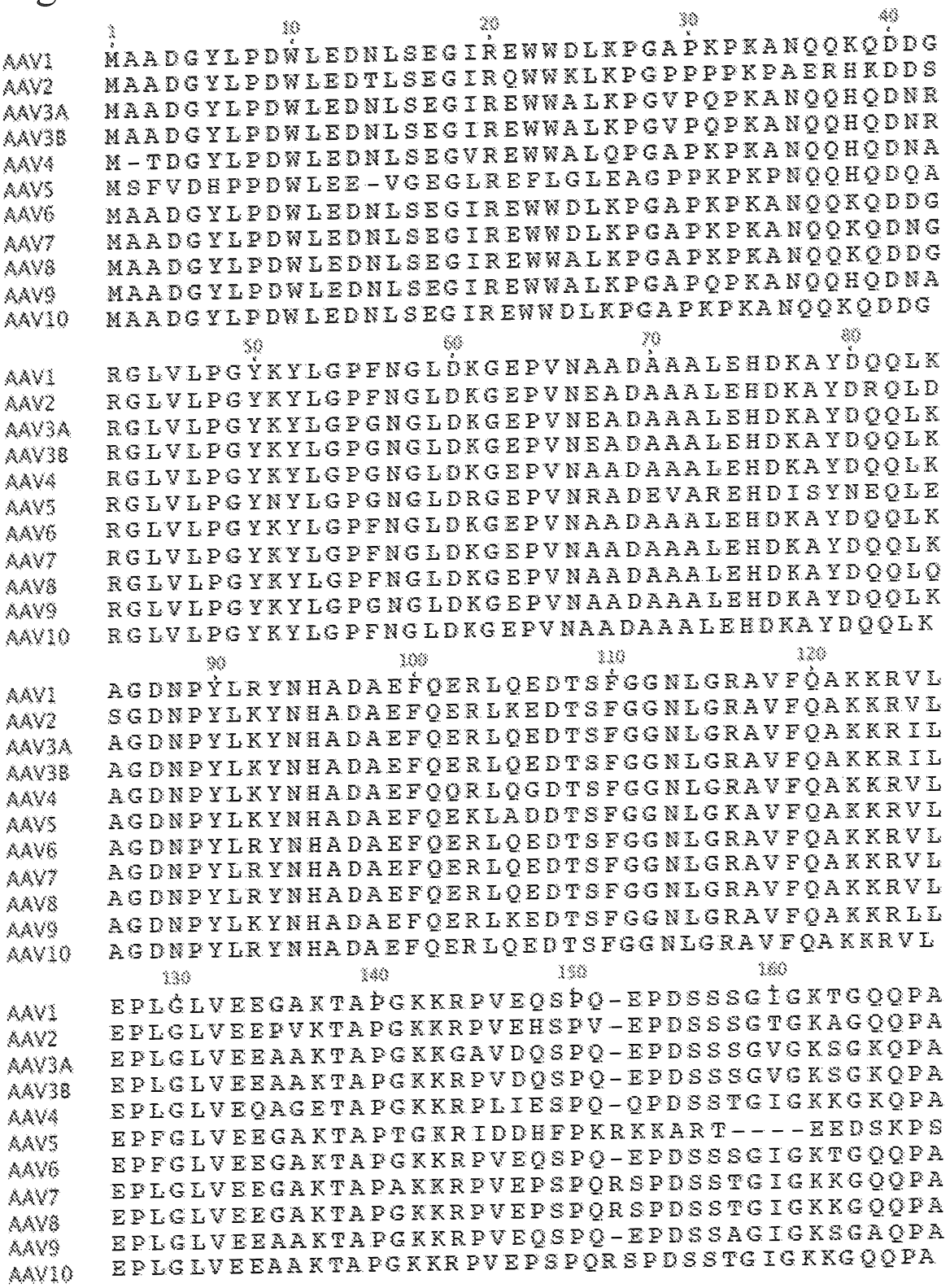

1 10 20 30 40
AAV1 MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKANQQKQDDG
AAV2 MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPPKPAERHKDDS
AAV3A MAADGYLPDWLEDNLSEGIREWWALKPGVPQPKANQQHQDNR
AAV3B MAADGYLPDWLEDNLSEGIREWWALKPGVPQPKANQQHQDNR
AAV4 M-TDGYLPDWLEDNLSEGVREWWALQPGAPKPKANQQHQDNA
AAV5 MSFVDHPPDWLEE-VGEGLREFLGLEAGPPKPKPNQQHQDQA
AAV6 MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKANQQKQDDG
AAV7 MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKANQQKQDNG
AAV8 MAADGYLPDWLEDNLSEGIREWWALKPGAPKPKANQQKQDDG
AAV9 MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNA
AAV10 MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKANQQKQDDG
50 60 70 80
AAV1 RGLVLPGYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLK
AAV2 RGLVLPGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLD
AAV3A RGLVLPGYKYLGPGNGLDKGEPVNEADAAALEHDKAYDQQLK
AAV3B RGLVLPGYKYLGPGNGLDKGEPVNEADAAALEHDKAYDQQLK
AAV4 RGLVLPGYKYLGPGNGLDKGEPVNAADAAALEHDKAYDQQLK
AAV5 RGLVLPGYNYLGPGNGLDRGEPVNRADEVAREHDISYNEQLE
AAV6 RGLVLPGYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLK
AAV7 RGLVLPGYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLK
AAV8 RGLVLPGYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLQ
AAV9 RGLVLPGYKYLGPGNGLDKGEPVNAADAAALEHDKAYDQQLK
AAV10 RGLVLPGYKYLGPFNGLDKGEPVNAADAAALEHDKAYDQQLK
90 100 110 120
AAV1 AGDNPYLRYNHADAEFQERLQEDTSFGGNLGRAVFQAKKRVL
AAV2 SGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQAKKRVL
AAV3A AGDNPYLKYNHADAEFQERLQEDTSFGGNLGRAVFQAKKRIL
AAV3B AGDNPYLKYNHADAEFQERLQEDTSFGGNLGRAVFQAKKRIL
AAV4 AGDNPYLKYNHADAEFQQRLQGDTSFGGNLGRAVFQAKKRVL
AAV5 AGDNPYLKYNHADAEFQEKLADDTSFGGNLGKAVFQAKKRVL
AAV6 AGDNPYLRYNHADAEFQERLQEDTSFGGNLGRAVFQAKKRVL
AAV7 AGDNPYLRYNHADAEFQERLQEDTSFGGNLGRAVFQAKKRVL
AAV8 AGDNPYLRYNHADAEFQERLQEDTSFGGNLGRAVFQAKKRLL
AAV9 AGDNPYLKYNHADAEFQERLKEDTSFGGNLGRAVFQAKKRVL
AAV10 AGDNPYLRYNHADAEFQERLQEDTSFGGNLGRAVFQAKKRVL
130 140 150 160
AAV1 EPLGLVEEGAKTAPGKKRPVEQSPQ-EPDSSSGIGKTGQQPA
AAV2 EPLGLVEEPVKTAPGKKRPVEHSPV-EPDSSSGTGKAGQQPA
AAV3A EPLGLVEEAAKTAPGKKGAVDQSPQ-EPDSSSGVGKSGKQPA
AAV3B EPLGLVEEAAKTAPGKKRPVDQSPQ-EPDSSSGVGKSGKQPA
AAV4 EPLGLVEQAGETAPGKKRPLIESPQ-QPDSSTGIGKKGKQPA
AAV5 EPFGLVEEGAKTAPTGKRIDDHFPKRKKART----EEDSKPS
AAV6 EPFGLVEEGAKTAPGKKRPVEQSPQ-EPDSSSGIGKTGQQPA
AAV7 EPLGLVEEGAKTAPAKKRPVEPSPQRSPDSSTGIGKKGQQPA
AAV8 EPLGLVEEGAKTAPGKKRPVEPSPQRSPDSSTGIGKKGQQPA
AAV9 EPLGLVEEAAKTAPGKKRPVEQSPQ-EPDSSAGIGKSGAQPA
AAV10 EPLGLVEEAAKTAPGKKRPVEPSPQRSPDSSTGIGKKGQQPA

Fig. 5 (tiếp)


170 180 190 200 210
AAV1 KKRLNFGQTGDSESVPDPQPLGEPPATP-AAVGPTTMASGGG
AAV2 RKRLNFGQTGDADSVPDPQPLGQPPAAP-SGLGTNTMATGSG
AAV3A RKRLNFGQTGDSESVPDPQPLGEPPAAP-TSLGSNTMASGGG
AAV3B RKRLNFGQTGDSESVPDPQPLGEPPAAP-TSLGSNTMASGGG
AAV4 KKKLVFE---DETGAGDGPPEGSTSGA--MSDDSEMRAAAGG
AAV5 T-------SSDAEAGPSGSQQLQIPAQPASSLGADTMSAGGG
AAV6 KKRLNFGQTGDSESVPDPQPLGEPPATP-AAVGPTTMASGGG
AAV7 RKRLNFGQTGDSESVPDPQPLGEPPAAP-SSVGSGTVAAGGG
AAV8 RKRLNFGQTGDSESVPDPQPLGEPPAAP-SGVGPNTMAAGGG
AAV9 KKRLNFGQTGDTESVPDPQPIGEPPAAP-SGVGSLTMASGGG
AAV10 KKRLNFGQTGESESVPDPQPIGEPPAGP-SGLGSGTMAAGGG
220 230 240 250
AAV1 APMADNNEGADDGVGNASGNWHCDSTWLGDRVITTSTRTWALP
AAV2 APMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALP
AAV3A APMADNNEGADGVGNSSGNWHCDSQWLGDRVITTSTRTWALP
AAV3B APMADNNEGADGVGNSSGNWHCDSQWLGDRVITTSTRTWALP
AAV4 AAV-EGGQGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLP
AAV5 GPLGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVLP
AAV6 APMADNNEGADGVGNASGNWHCDSTWLGDRVITTSTRTWALP
AAV7 APMADNNEGADGVGNASGNWHCDSTWLGDRVITTSTRTWALP
AAV8 APMADNNEGADGVGSSSGNWHCDSTWLGDRVITTSTRTWALP
AAV9 APVADNNEGADGVGSSSGNWHCDSQWLGDRVITTSTRTWALP
AAV10 APMADNNEGADGVGSSSGNWHCDSTWLGDRVITTSTRTWALP
260 270 280 290
AAV1 TYNNHLYKQISSA-STGASNDNHYFGYSTPWGYFDFNRFHCH
AAV2 TYNNHLYKQISSQ--SGASNDNHYFGYSTPWGYFDFNRFHCH
AAV3A TYNNHLYKQISSQ--SGASNDNHYFGYSTPWGYFDFNRFHCH
AAV3B TYNNHLYKQISSQ--SGASNDNHYFGYSTPWGYFDFNRFHCH
AAV4 TYNNHLYKRLGE-----SLQSNTYNGFSTPWGYFDFNRFHCH
AAV5 SYNNHQYREIKSGSVDG-SNANAYFGYSTPWGYFDFNRFHSH
AAV6 TYNNHLYKQISSA-STGASNDNHYFGYSTPWGYFDFNRFHCH
AAV7 TYNNHLYKQISSE-TAGSTNDNTYFGYSTPWGYFDFNRFHCH
AAV8 TYNNHLYKQISNGTSGGATNDNTYFGYSTPWGYFDFNRFHCH
AAV9 TYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCH
AAV10 TYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDFNRFHCH
300 310 320 330
AAV1 FSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTTNDGVTTI
AAV2 FSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTI
AAV3A FSPRDWQRLINNNWGFRPKKLSFKLFNIQVRGVTQNDGTTTI
AAV3B FSPRDWQRLINNNWGFRPKKLSFKLFNIQVKEVTQNDGTTTI
AAV4 FSPRDWQRLINNNWGMRPKAMRVKIFNIQVKEVTTSNGETTV
AAV5 WSPRDWQRLINNYWGFRPRSLRVKIFNIQVKEVTVQDSTTTI
AAV6 FSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTTNDGVTTI
AAV7 FSPRDWQRLINNNWGFRPKKLRFKLFNIQVKEVTTNDGVTTI
AAV8 FSPRDWQRLINNNWGFRPKRLSFKLFNIQVKEVTQNEGTKTI
AAV9 FSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTI
AAV10 FSPRDWQRLINNNWGFRPKRLSFKLFNIQVKEVTQNEGTKTI

Fig. 5 (tiếp)


340 350 360 370
AAV1 ANNLTSTVQVFSDSEYQLPYVLGSAHQGCLPPFPADVFMIPQ
AAV2 ANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQ
AAV3A ANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQ
AAV3B ANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQ
AAV4 ANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVPQ
AAV5 ANNLTSTVQVFTDDDYQLPYVVGNGTEGCLPAFPPQVFTLPQ
AAV6 ANNLTSTVQVFSDSEYQLPYVLGSAHQGCLPPFPADVFMIPQ
AAV7 ANNLTSTIQVFSDSEYQLPYVLGSAHQGCLPPFPADVFMIPQ
AAV8 ANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQ
AAV9 ANNLTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQ
AAV10 ANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQ
380 390 400 410 420
AAV1 YGYLTLN--NGS-QAVGRSSFYCLEYFPSQMLRTGNNFTFSY
AAV2 YGYLTLN--NGS-QAVGRSSFYCLEYFPSQMLRTGNNFTFSY
AAV3A YGYLTLN--NGS-QAVGRSSFYCLEYFPSQMLRTGNNFQFSY
AAV3B YGYLTLN--NGS-QAVGRSSFYCLEYFPSQMLRTGNNFQFSY
AAV4 YGYCGLVTGNTSQQQTDRNAFYCLEYFPSQMLRTGNNFEITY
AAV5 YGYATLNRDNTE-NPTERSSFFCLEYFPSKMLRTGNNFEFTY
AAV6 YGYLTLN--NGS-QAVGRSSFYCLEYFPSQMLRTGNNFTFSY
AAV7 YGYLTLN--NGS-QSVGRSSFYCLEYFPSQMLRTGNNFEFSY
AAV8 YGYLTLN--NGS-QAVGRSSFYCLEYFPSQMLRTGNNFQFTY
AAV9 YGYLTLN--DGS-QAVGRSSFYCLEYFPSQMLRTGNNFQFSY
AAV10 YGYLTLN--NGS-QAVGRSSFYCLEYFPSQMLRTGNNFEFSY
430 440 450 460
AAV1 TFEEVPFHSSYAHSQSLDRLMNPLIDQYLYYLNRTQNQ-SGS
AAV2 TFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRT-NTPSGT
AAV3A TFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLNRTQGTTSGT
AAV3B TFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLNRTQGTTSGT
AAV4 SFEKVPFHSMYAHSQSLDRLMNPLIDQYLWGLQSTTTGTTLN
AAV5 NFEEVPFHSSFAPSQNLFKLANPLVDQYLYRFVSTNNTGG--
AAV6 TFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLNRTQNQ-SGS
AAV7 SFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLARTQSNPGGT
AAV8 TFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTQTT-GGT
AAV9 EFENVPFHSSYAHSQSLDRLMNPLIDQYLYYLSKTIN--GSG
AAV10 TFEDVPFHSSYAHSQSLDRLMNPLIDQYLYYLSRTQST-GGT
470 480 490 500
AAV1 AQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDN-
AAV2 TTQSRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADN-
AAV3A TNQSRLLFSQAGPQSMSLQARNWLPGPCYRQQRLSKTANDN-
AAV3B TNQSRLLFSQAGPQSMSLQARNWLPGPCYRQQRLSKTANDN-
AAV4 AGTATTNFTKLRPTNFSNFKKNWLPGPSIKQQGFSKTANQNY
AAV5 -----VQFNKNLAGRYANTYKNWFPGPMGRTQGWNLGSGVNR
AAV6 AQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDN-
AAV7 AGNRELQFYQGGPSTMAEQAKNWLPGPCFRQQRVSKTLDQN-
AAV8 ANTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQN-
AAV9 QNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQN-
AAV10 QGTQQLLFSQAGPANMSAQAKNWLPGPCYRQQRVSTTLSQN-

Fig. 5 (tiếp)

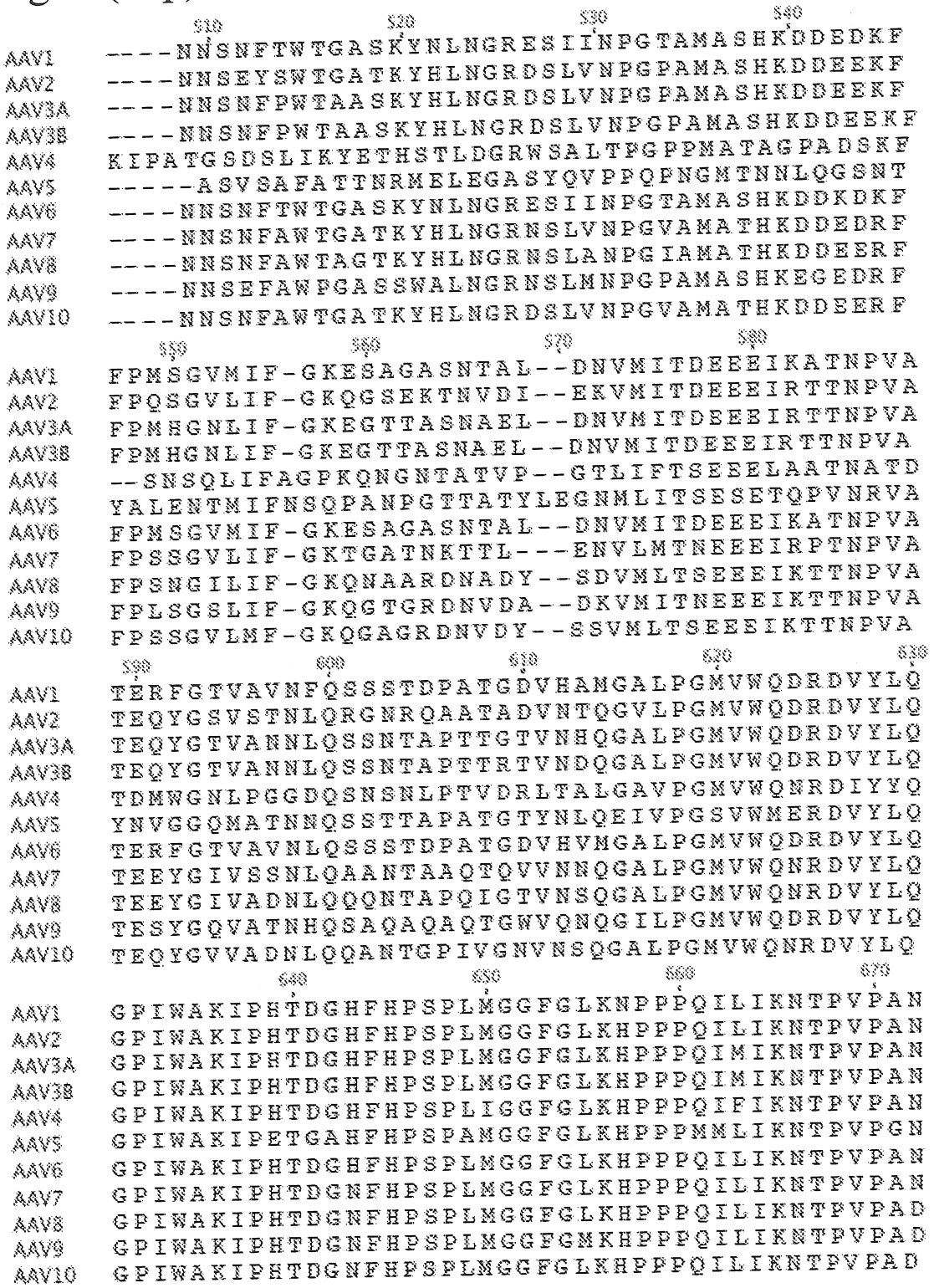

510 520 530 540
AAV1 ----NNSNFTWTGASKYNLNGRESIINPGTAMASHKDDEDKF
AAV2 ----NNSEYSWTGATKYHLNGRDSLVNPGPAMASHKDDEEKF
AAV3A ----NNSNFPWTAASKYHLNGRDSLVNPGPAMASHKDDEEKF
AAV3B ----NNSNFPWTAASKYHLNGRDSLVNPGPAMASHKDDEEKF
AAV4 KIPATGSDSLIKYETHSTLDGRWSALTPGPPMATAGPADSKF
AAV5 -----ASVSAFATTNRMELEGASYQVPPQPNGMTNNLQGSNT
AAV6 ----NNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDKF
AAV7 ----NNSNFAWTGATKYHLNGRNSLVNPGVAMATHKDDEDRF
AAV8 ----NNSNFAWTAGTKYHLNGRNSLANPGIAMATHKDDEERF
AAV9 ----NNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRF
AAV10 ----NNSNFAWTGATKYHLNGRDSLVNPGVAMATHKDDEERF
550 560 570 580
AAV1 FPMSGVMIF-GKESAGASNTAL--DNVMITDEEEIKATNPVA
AAV2 FPQSGVLIF-GKQGSEKTNVDI--EKVMITDEEEIRTTNPVA
AAV3A FPMHGNLIF-GKEGTTASNAEL--DNVMITDEEEIRTTNPVA
AAV3B FPMHGNLIF-GKEGTTASNAEL--DNVMITDEEEIRTTNPVA
AAV4 --SNSQLIFAGPKQNGNTATVP--GTLIFTSEEELAATNATD
AAV5 YALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVA
AAV6 FPMSGVMIF-GKESAGASNTAL--DNVMITDEEEIKATNPVA
AAV7 FPSSGVLIF-GKTGATNKTTL---ENVLMTNEEEIRPTNPVA
AAV8 FPSNGILIF-GKQNAARDNADY--SDVMLTSEEEIKTTNPVA
AAV9 FPLSGSLIF-GKQGTGRDNVDA--DKVMITNEEEIKTTNPVA
AAV10 FPSSGVLMF-GKQGAGRDNVDY--SSVMLTSEEEIKTTNPVA
590 600 610 620 630
AAV1 TERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDRDVYLQ
AAV2 TEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVYLQ
AAV3A TEQYGTVANNLQSSNTAPTTGTVNHQGALPGMVWQDRDVYLQ
AAV3B TEQYGTVANNLQSSNTAPTTRTVNDQGALPGMVWQDRDVYLQ
AAV4 TDMWGNLPGGDQSNSNLPTVDRLTALGAVPGMVWQNRDIYYQ
AAV5 YNVGGQMATNNQSSTTAPATGTYNLQEIVPGSVWMERDVYLQ
AAV6 TERFGTVAVNLQSSSTDPATGDVHVMGALPGMVWQDRDVYLQ
AAV7 TEEYGIVSSNLQAANTAAQTQVVNNQGALPGMVWQNRDVYLQ
AAV8 TEEYGIVADNLQQQNTAPQIGTVNSQGALPGMVWQNRDVYLQ
AAV9 TESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDRDVYLQ
AAV10 TEQYGVVADNLQQANTGPIVGNVNSQGALPGMVWQNRDVYLQ
640 650 660 670
AAV1 GPIWAKIPHTDGHFHPSPLMGGFGLKNPPPQILIKNTPVPAN
AAV2 GPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPAN
AAV3A GPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQIMIKNTPVPAN
AAV3B GPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQIMIKNTPVPAN
AAV4 GPIWAKIPHTDGHFHPSPLIGGFGLKHPPPQIFIKNTPVPAN
AAV5 GPIWAKIPETGAHFHPSPAMGGFGLKHPPPMMLIKNTPVPGN
AAV6 GPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPAN
AAV7 GPIWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTPVPAD
AAV8 GPIWAKIPHTDGNFHPSPLMGGFGMKHPPPQILIKNTPVPAD
AAV9 GPIWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTPVPAD
AAV10 GPIWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTPVPAD

Fig. 5 (tiếp)

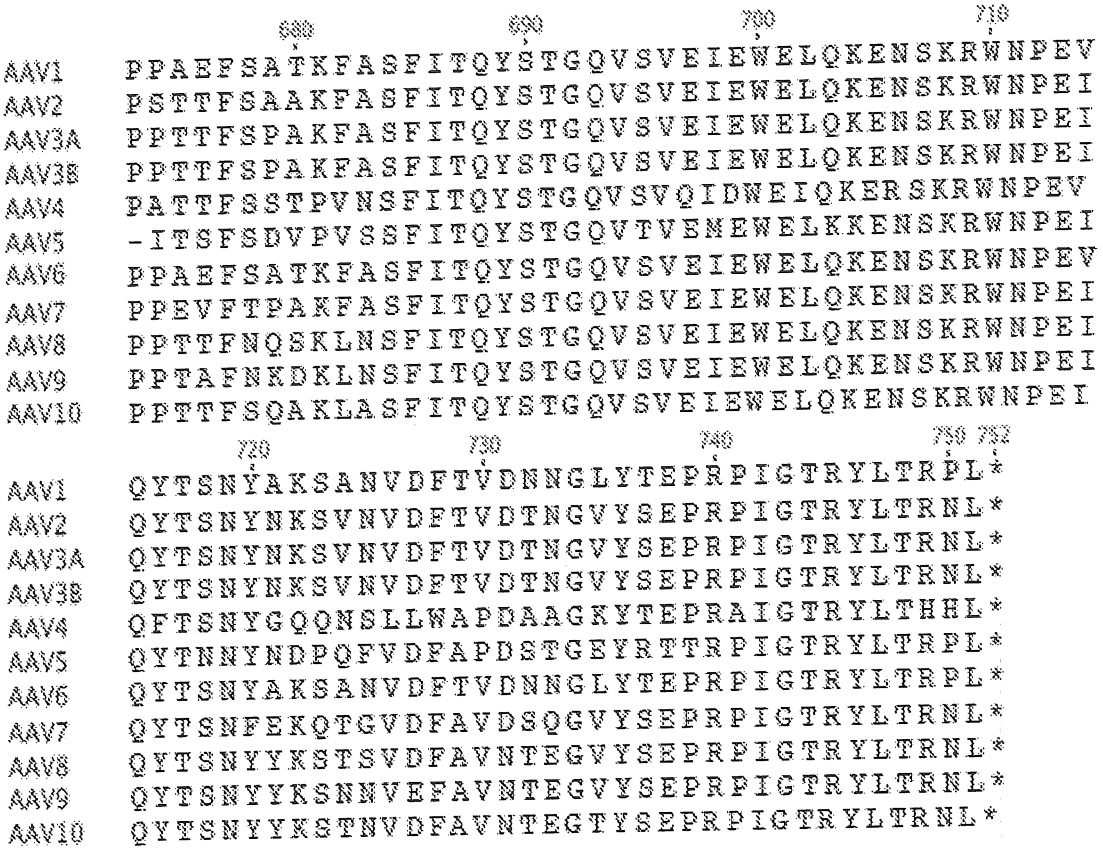

680 690 700 710
AAV1 PPAEFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEV
AAV2 PSTTFSAAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEI
AAV3A PPTTFSPAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEI
AAV3B PPTTFSPAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEI
AAV4 PATTFSSTPVNSFITQYSTGQVSVQIDWEIQKERSKRWNPEV
AAV5 -ITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEI
AAV6 PPAEFSATKFASFITQYSTGQVSVEIEWELQKENSKRWNPEV
AAV7 PPEVFTPAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEI
AAV8 PPTTFNQSKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEI
AAV9 PPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEI
AAV10 PPTTFSQAKLASFITQYSTGQVSVEIEWELQKENSKRWNPEI
720 730 740 750 752
AAV1 QYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL*
AAV2 QYTSNYNKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL*
AAV3A QYTSNYNKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL*
AAV3B QYTSNYNKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL*
AAV4 QFTSNYGQQNSLLWAPDAAGKYTEPRAIGTRYLTHHL*
AAV5 QYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL*
AAV6 QYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL*
AAV7 QYTSNFEKQTGVDFAVDSQGVYSEPRPIGTRYLTRNL*
AAV8 QYTSNYYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL*
AAV9 QYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL*
AAV10 QYTSNYYKSTNVDFAVNTEGTYSEPRPIGTRYLTRNL*

Fig. 6A

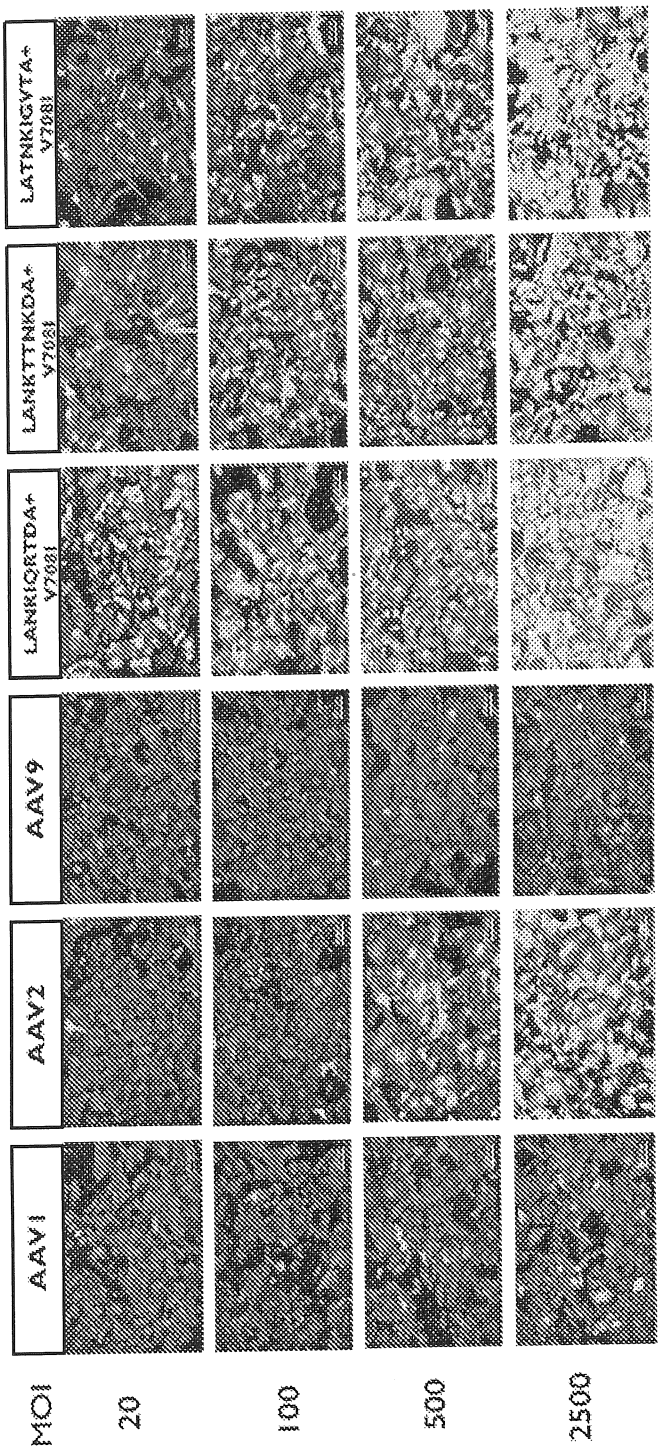
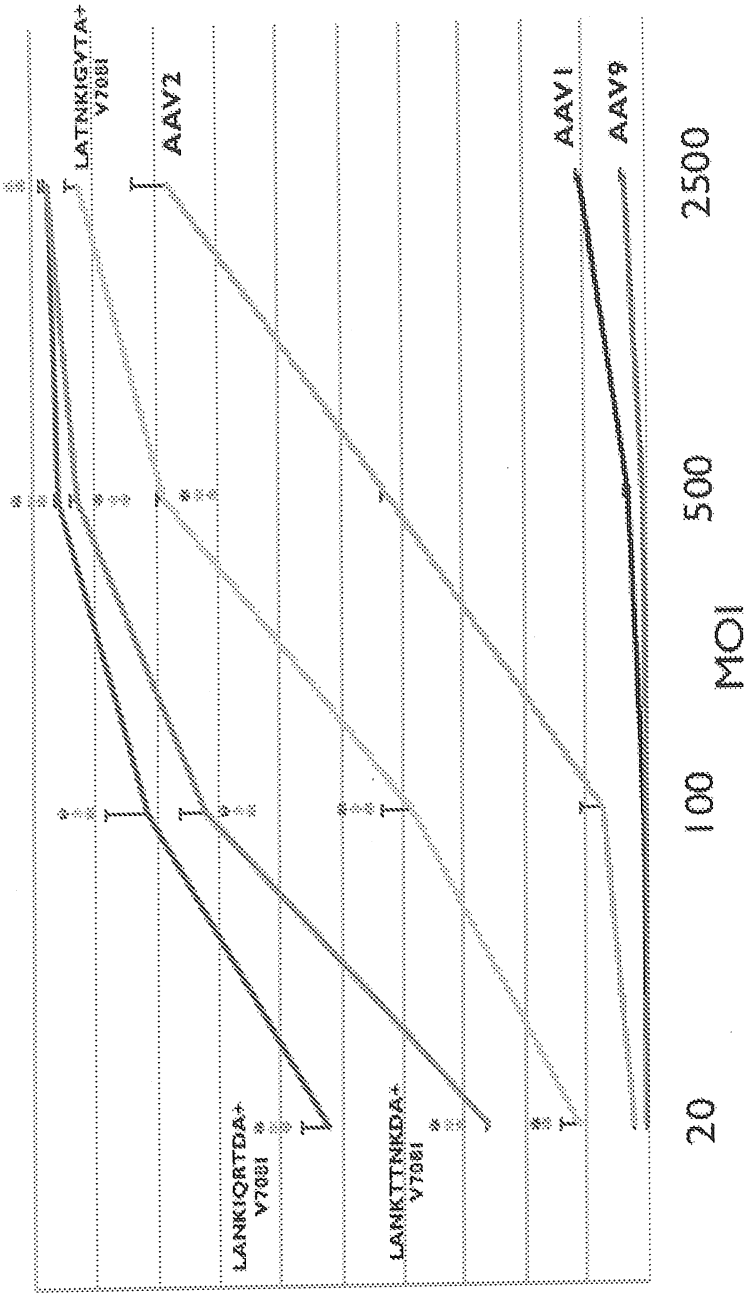

eTnTesGFP/DAPI
AAV1
AAV2
AAV9
LANRIQRTDA+ V7081
LANKTTNKDA+ V7081
LATNKIGVTA+ V7081
MOI
20
100
500
2500
Thời gian Pi: 6 ngày

Fig. 6B



n=3
trung bình ± độ lệch chuẩn
Thử nghiệm T hai đầu
* $p<0,05$

Fig. 6C

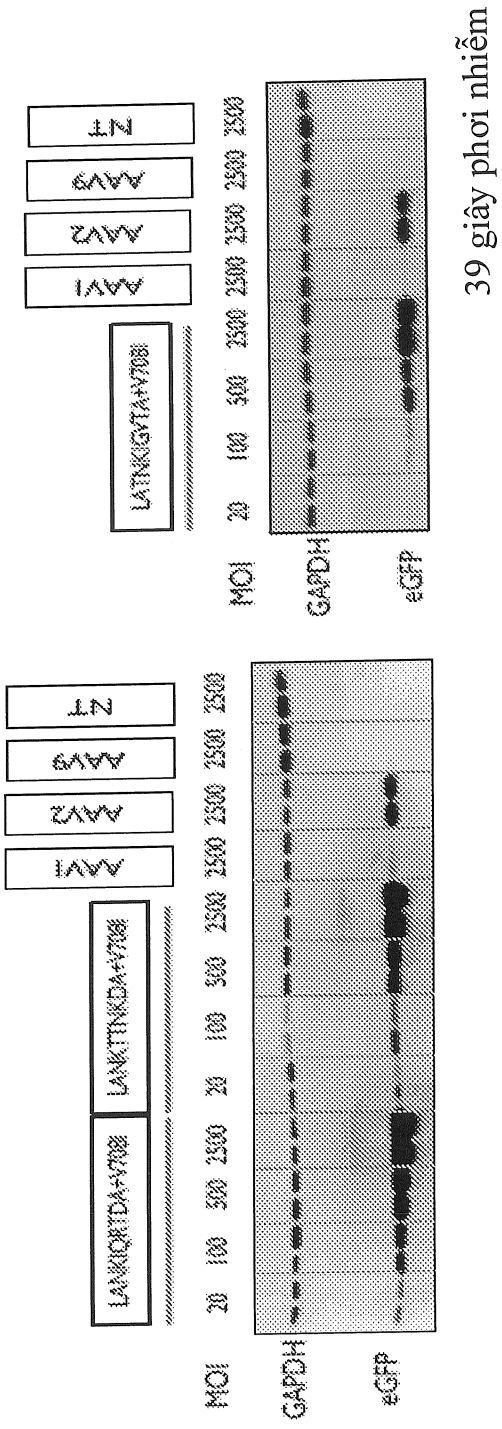

LAKKQRTDA+V708
LANKTTNKDA+V708
N7
A4V2
A4V9
N7
MOI
20
100
500
2500
GAPDH
eGFP
39 giây phoi nhiễm

Fig. 6D

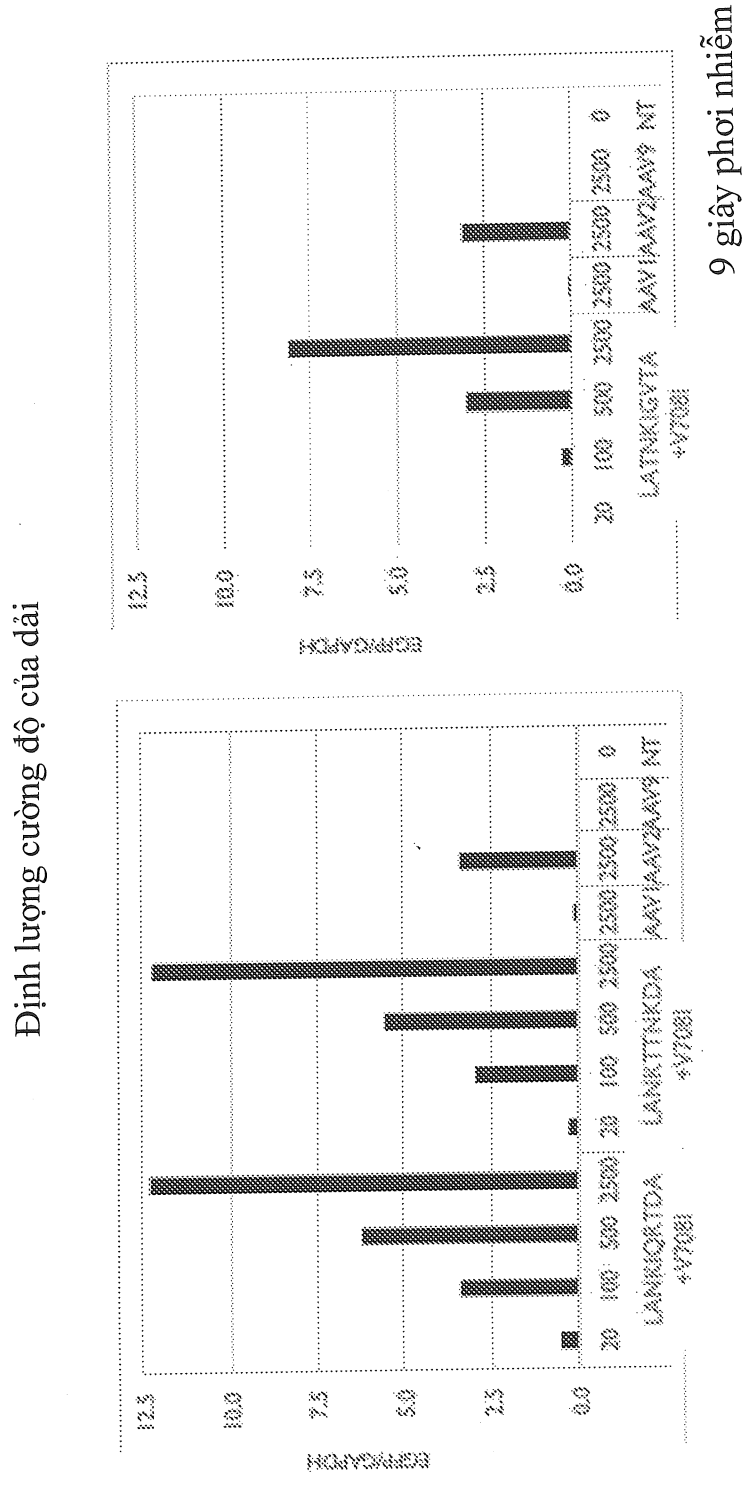

Định lượng cường độ của dải
9 giây phơi nhiễm
12.3
10.0
7.5
5.0
2.5
0.0
20 100 500 2500
LANMQRTDA
*Y708I
20 100 500 2500
LANETTNKDA
*Y708I
2500 2500 2500 0
AAV1AAV2AAV9 NT
2500 2500 2500 0
AAV1AAV2AAV9 NT

Fig. 6E

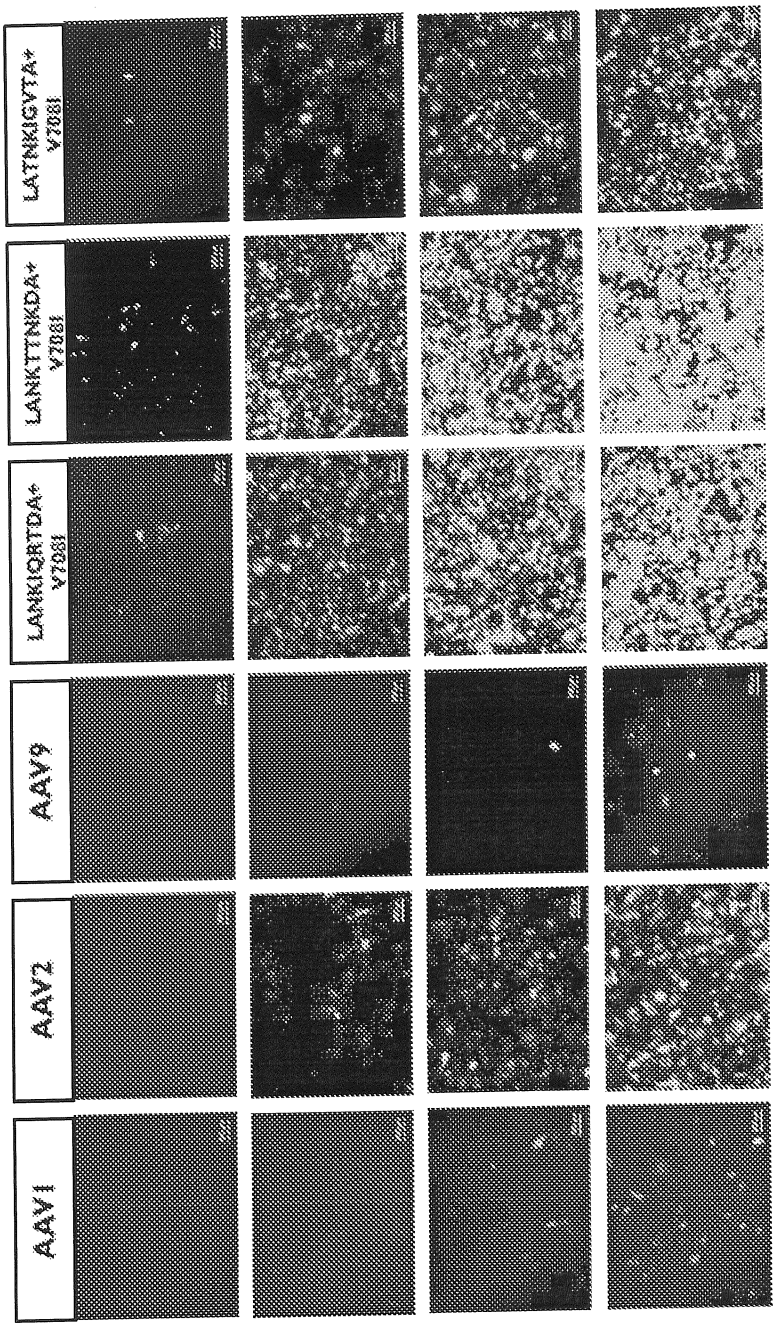
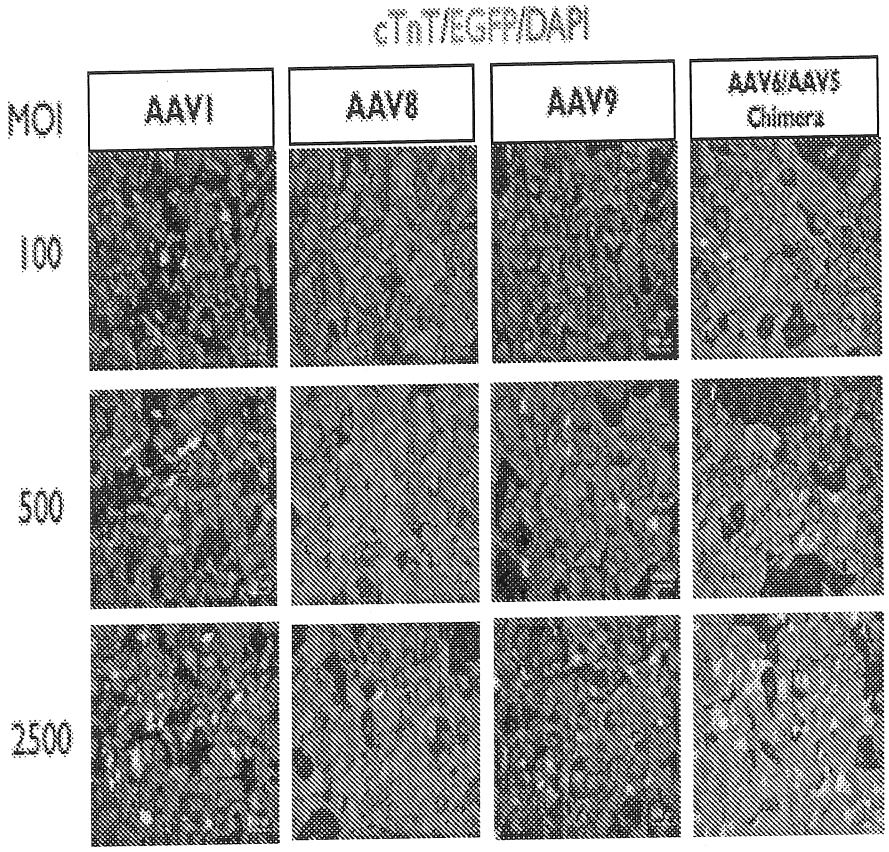
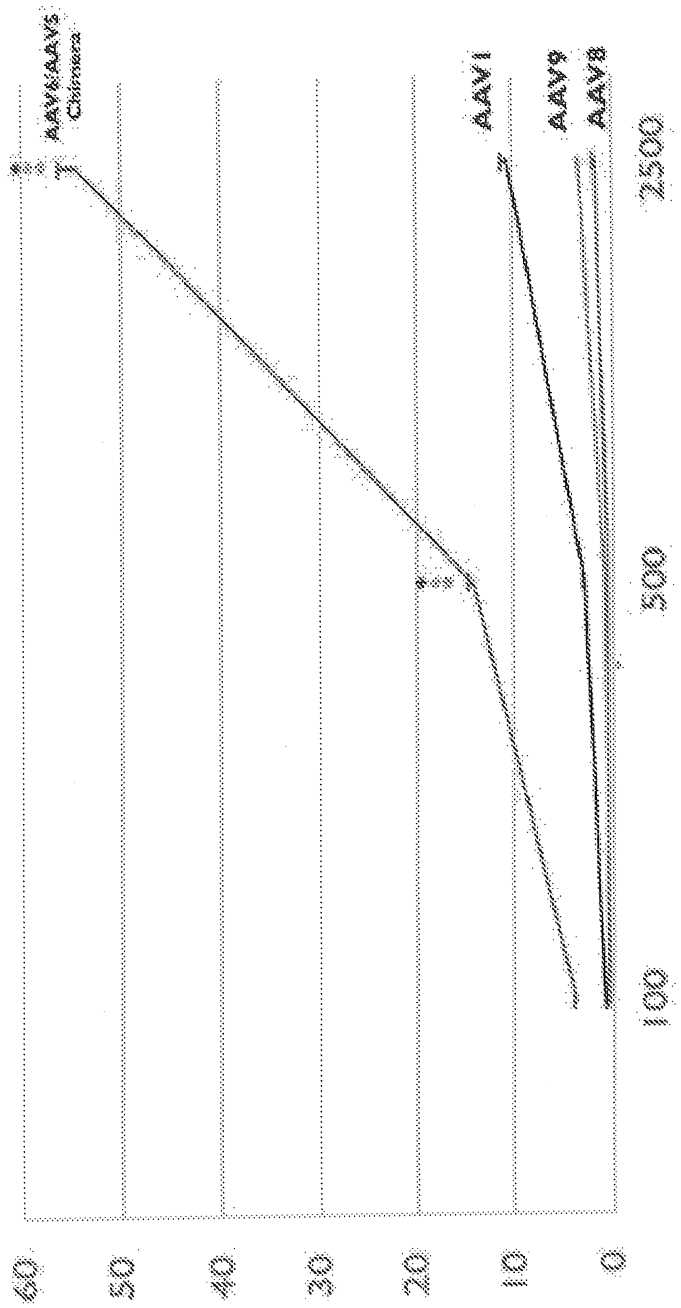

EGFP
AAV1
AAV2
AAV9
LANKIQRTDA+ V708I
LANKTTNKDA+ V708I
LATNKIGVTA+ V708I
Ngày
1
2
3
5
MOI: 500

Fig. 7A



Thời gian Pi: 6 ngày

Fig. 7B



n=3
trung bình ± độ lệch chuẩn
Thử nghiệm T hai đầu
* $p<0,05$

Fig. 7C

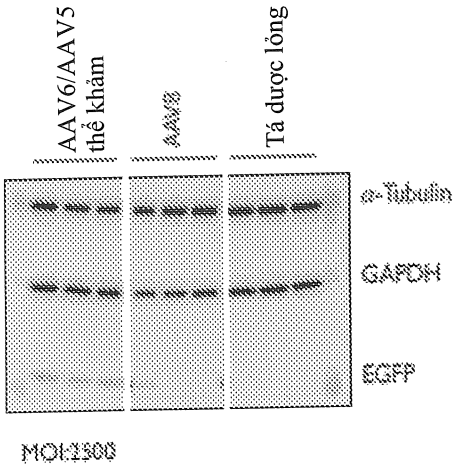


Fig. 7D

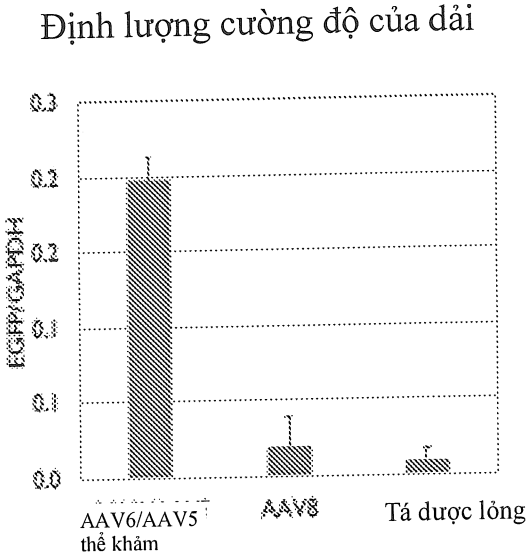


Fig. 7E

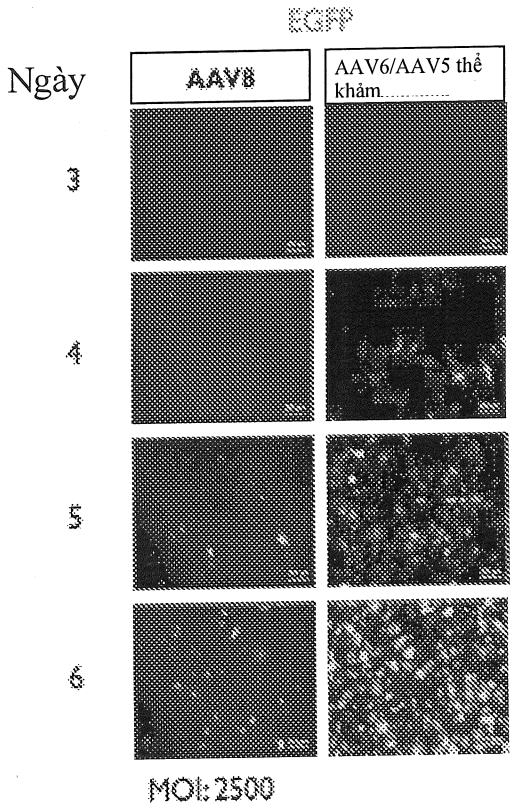


Fig. 8A

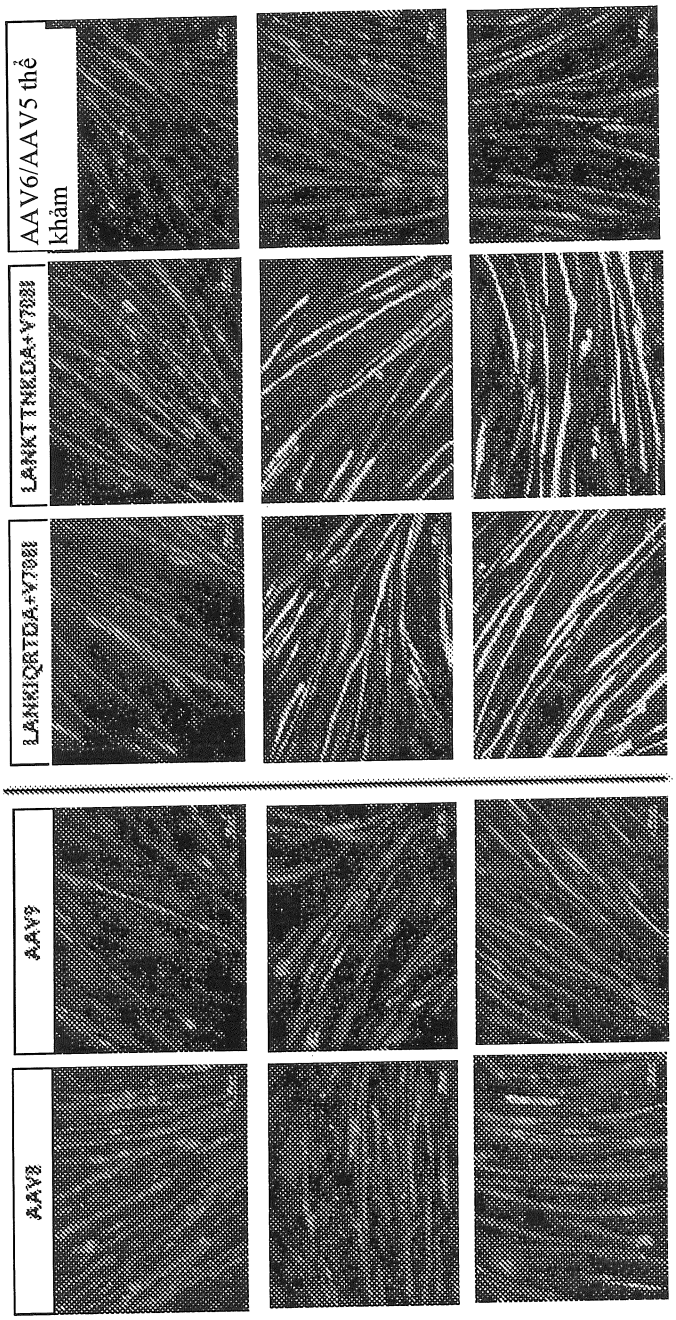
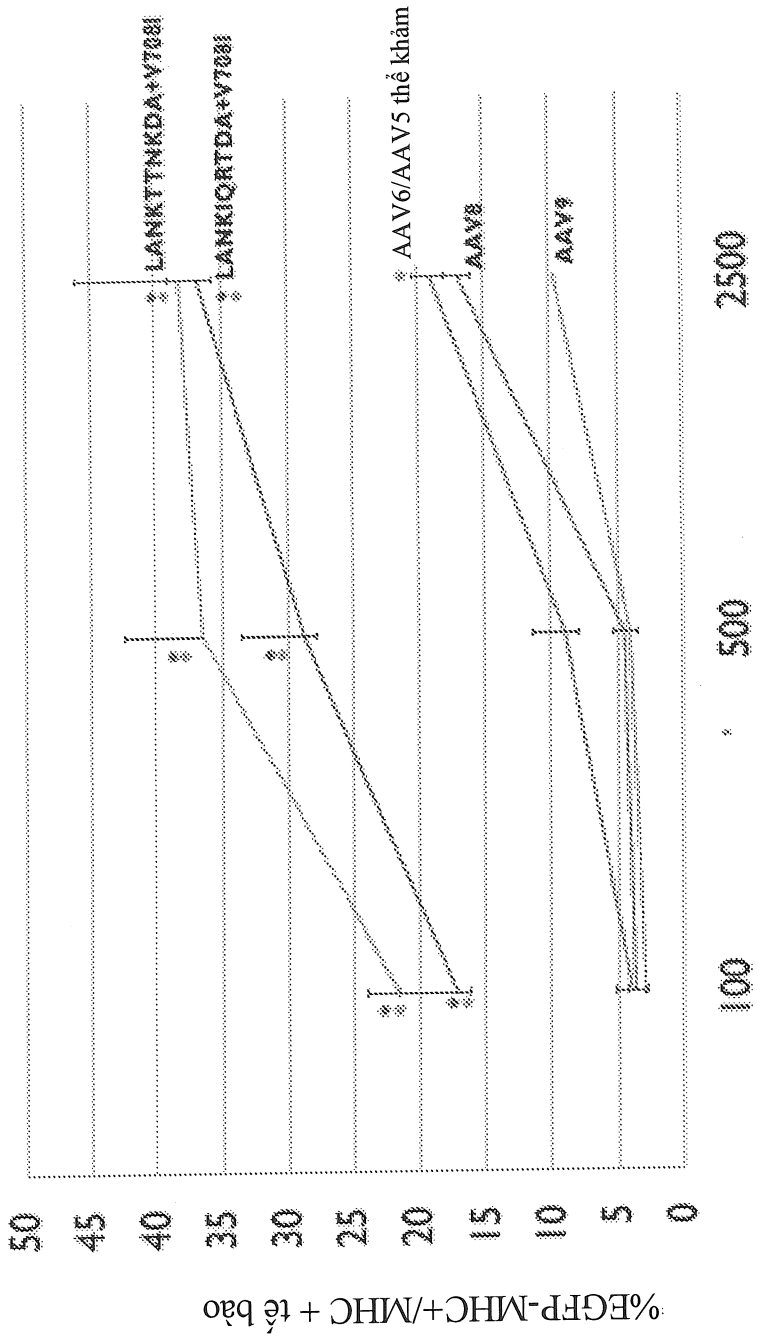

Chuỗi nặng myosin (MHC)/AAV-EGDP/DAPI
MOI
AAV8
AAV9
LANKTTMKDA+Y730F
LANKTTMKDA+Y730F
AAV6/AAV5 thể khảm
100
500
2500
Thời gian Pi: 7 ngày

Fig. 8B



n=3
trung bình ± độ lệch chuẩn
Thử nghiệm T hai đầu
* $p<0,05$

Fig. 8C

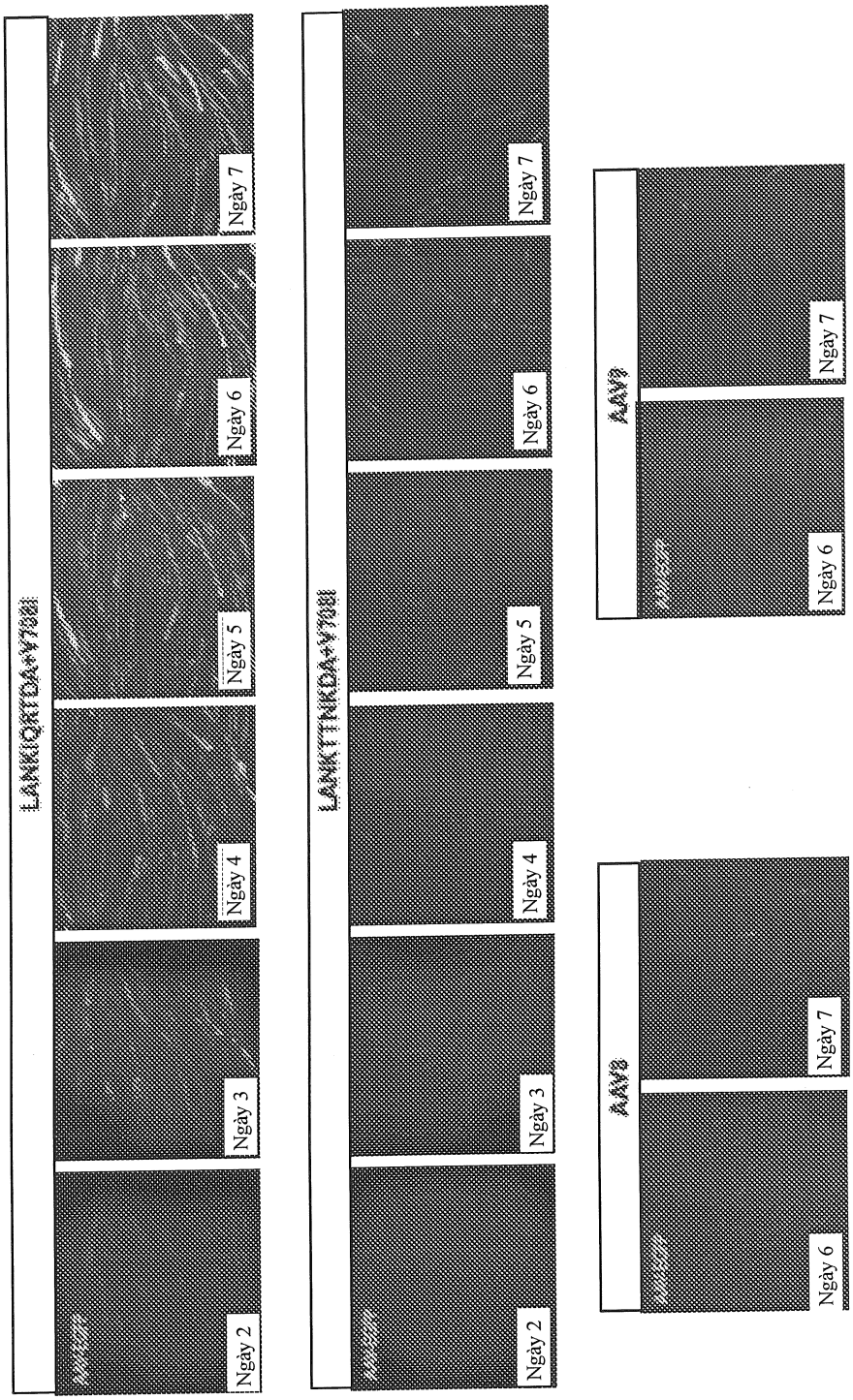

LANKIQRTDA+VT5d
Ngày 2
Ngày 3
Ngày 4
Ngày 5
Ngày 6
Ngày 7
LANKTTNRDA+VT5d
Ngày 2
Ngày 3
Ngày 4
Ngày 5
Ngày 6
Ngày 7
AAV8
Ngày 6
Ngày 7
AAV8
Ngày 6
Ngày 7

Fig. 9A

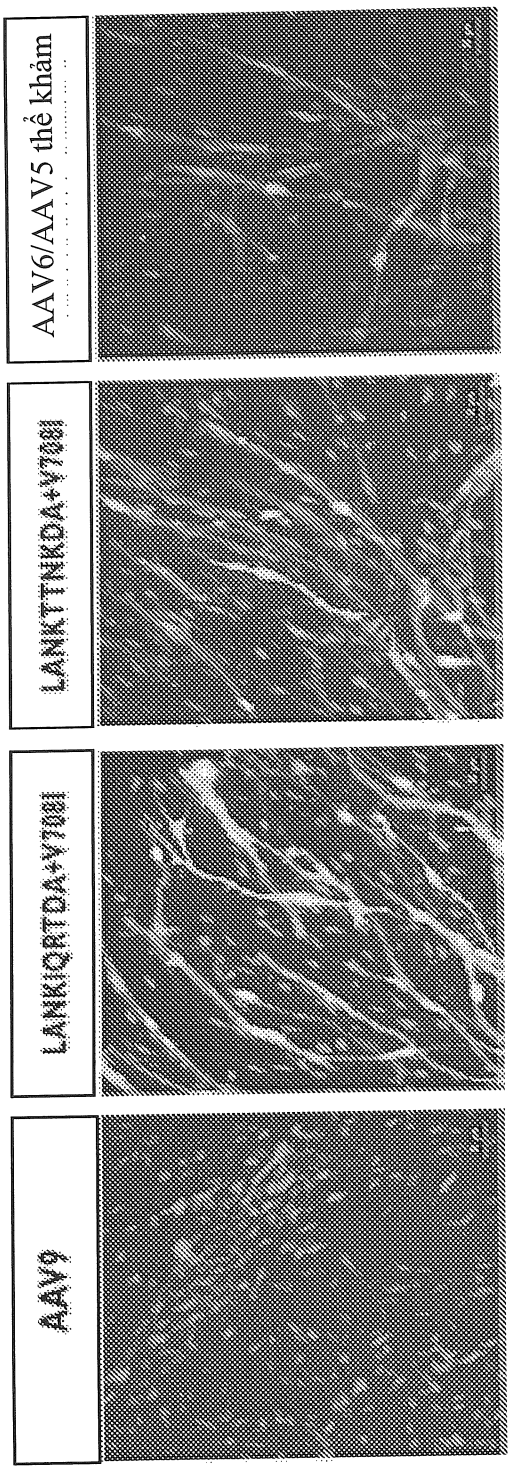
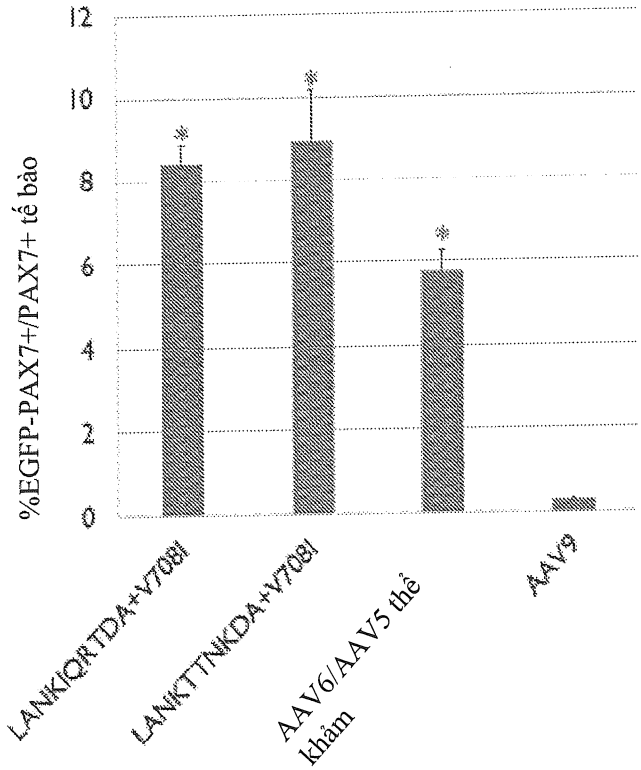

PAX7/EGFP/DAPI
AAV9
LANKIQRTDA+Y708I
LANKTTNKDA+Y708I
AAV6/AAV5 thể khảm
Thời gian PI: 6 ngày
MOI: 500

Fig.9B



n=3
trung bình ± độ lệch chuẩn
thử nghiệm T hai đầu
* p<0,05

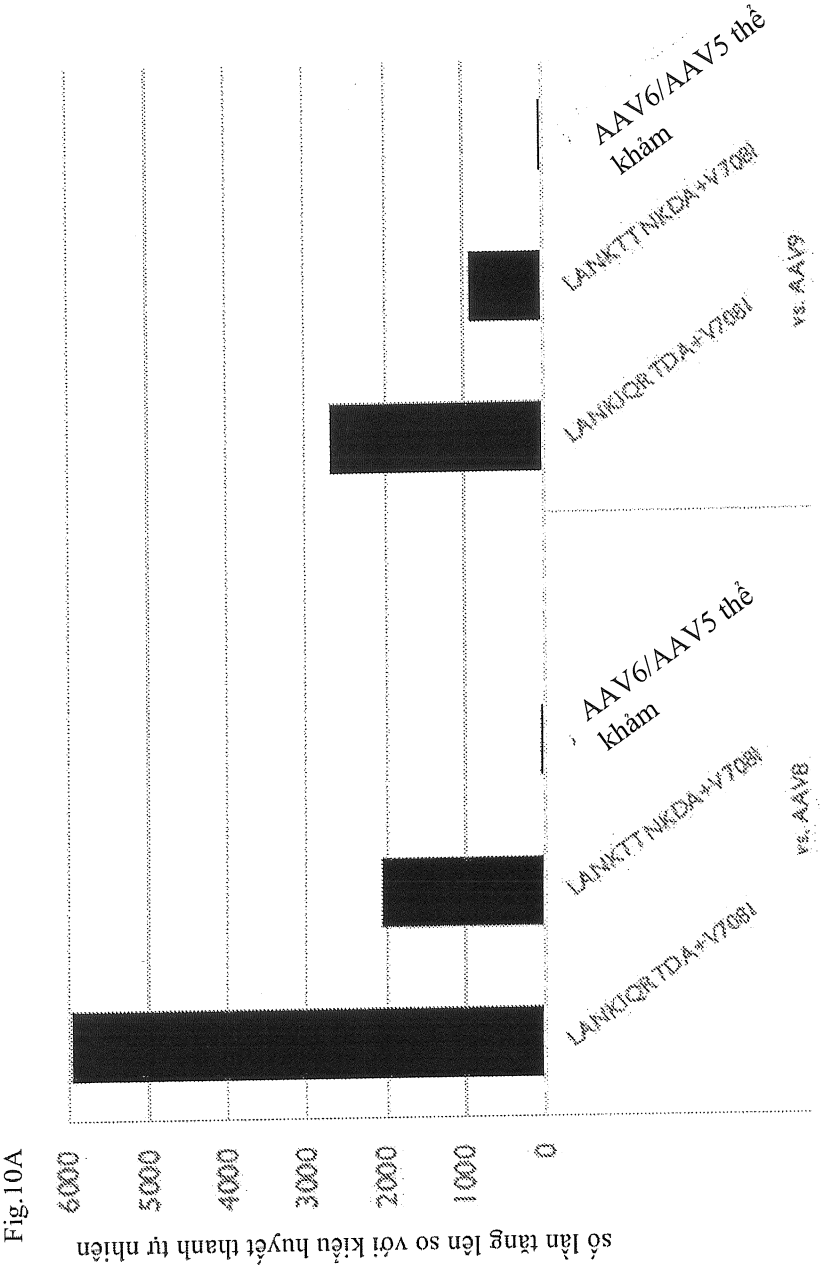

AAV6/AAV5 thể khảm
LANKTTNKDA+V708I
LANKTQRTDA+V708I
vs. AAV5
AAV6/AAV5 thể khảm
LANKTTNKDA+V708I
LANKTQRTDA+V708I
vs. AAV6
6000
5000
4000
3000
2000
1000
0
số lần tăng lên so với kiểu huyết thanh tự nhiên

Fig.10A

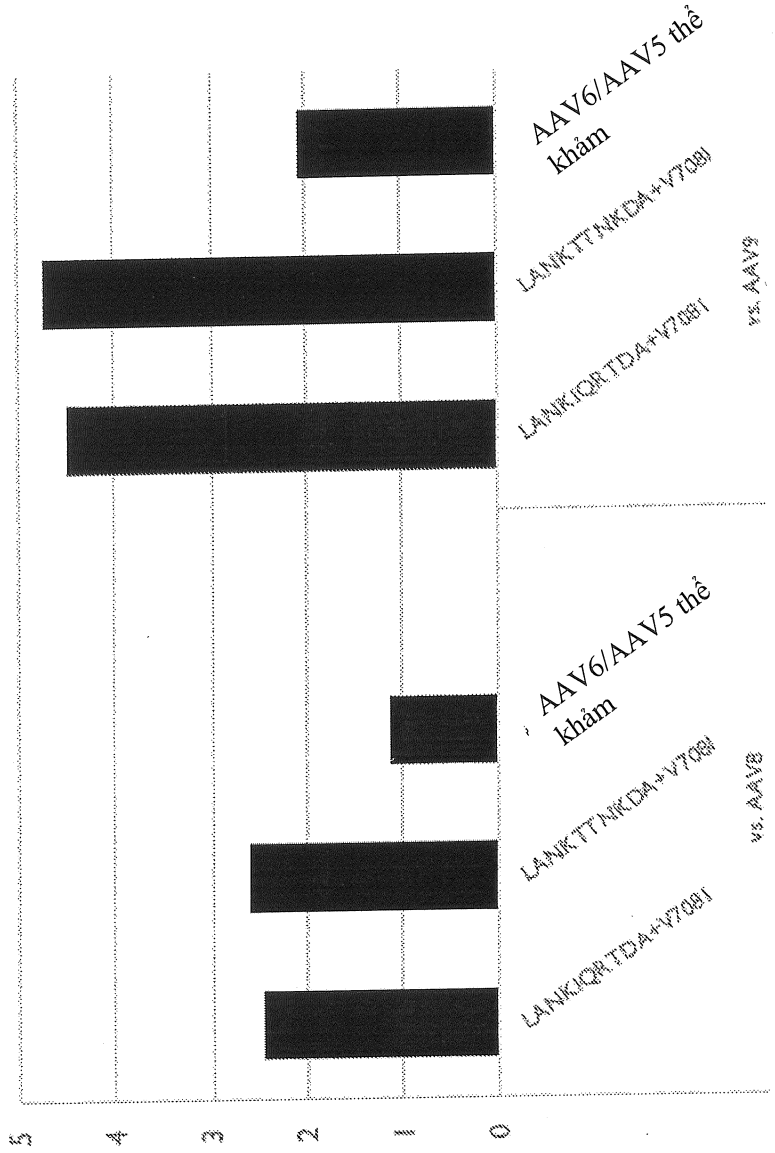

AAV6/AAV5 thể khảm
LANKTTNKCDA+V708I
LANKIQRTCDA+V708I
vs. AAV9
AAV6/AAV5 thể khảm
LANKTTNKCDA+V708I
LANKIQRTCDA+V708I
vs. AAV8
0
1
2
3
4
5
số lần tăng lên so với kiểu huyết thanh tự nhiên

Fig. 10B

Fig. 11A

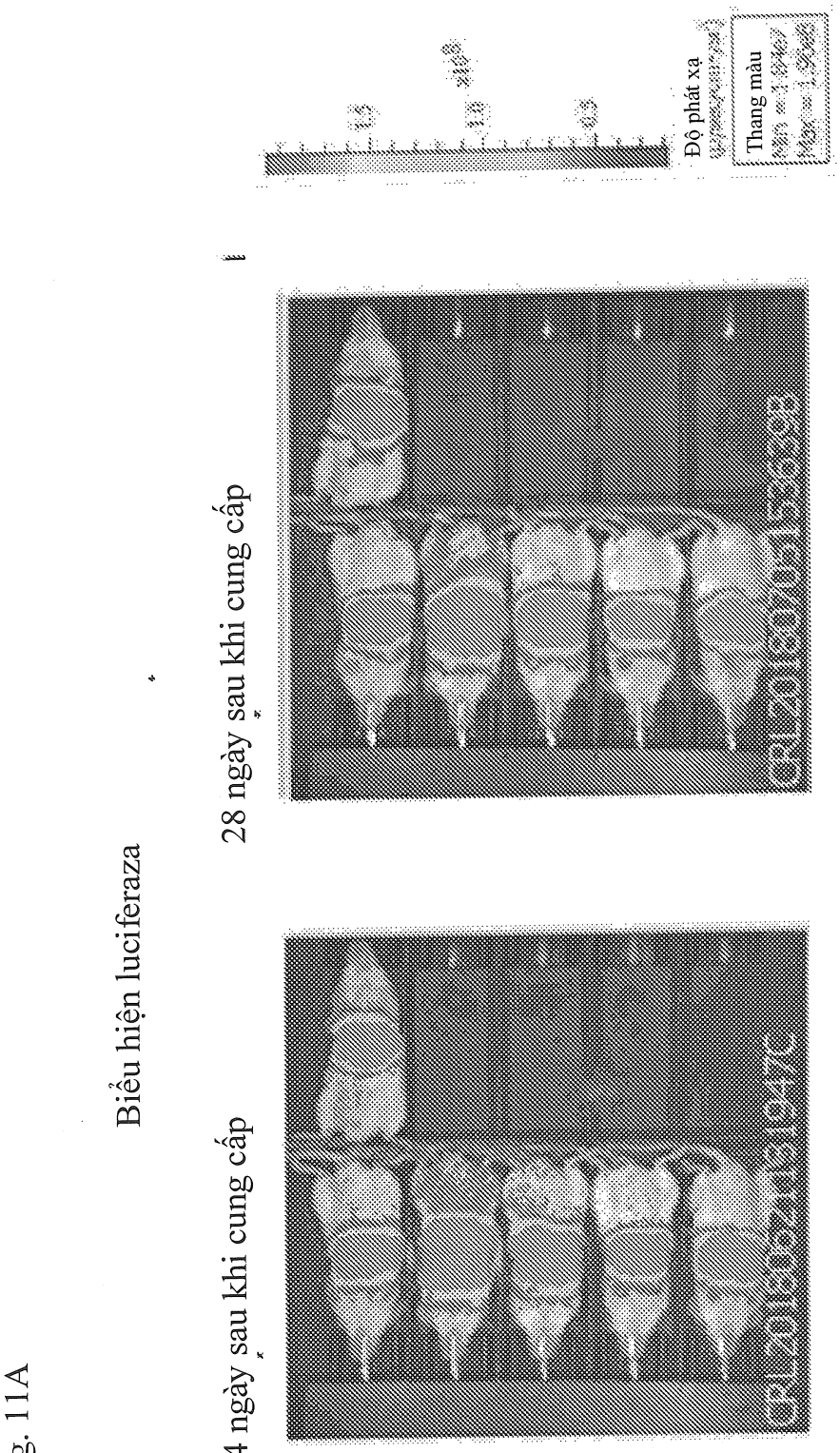

Biểu hiện luciferaza
14 ngày sau khi cung cấp
28 ngày sau khi cung cấp
Độ phát xạ
Thang màu

Fig. 11B

Biểu hiện luciferaza trong từng mô

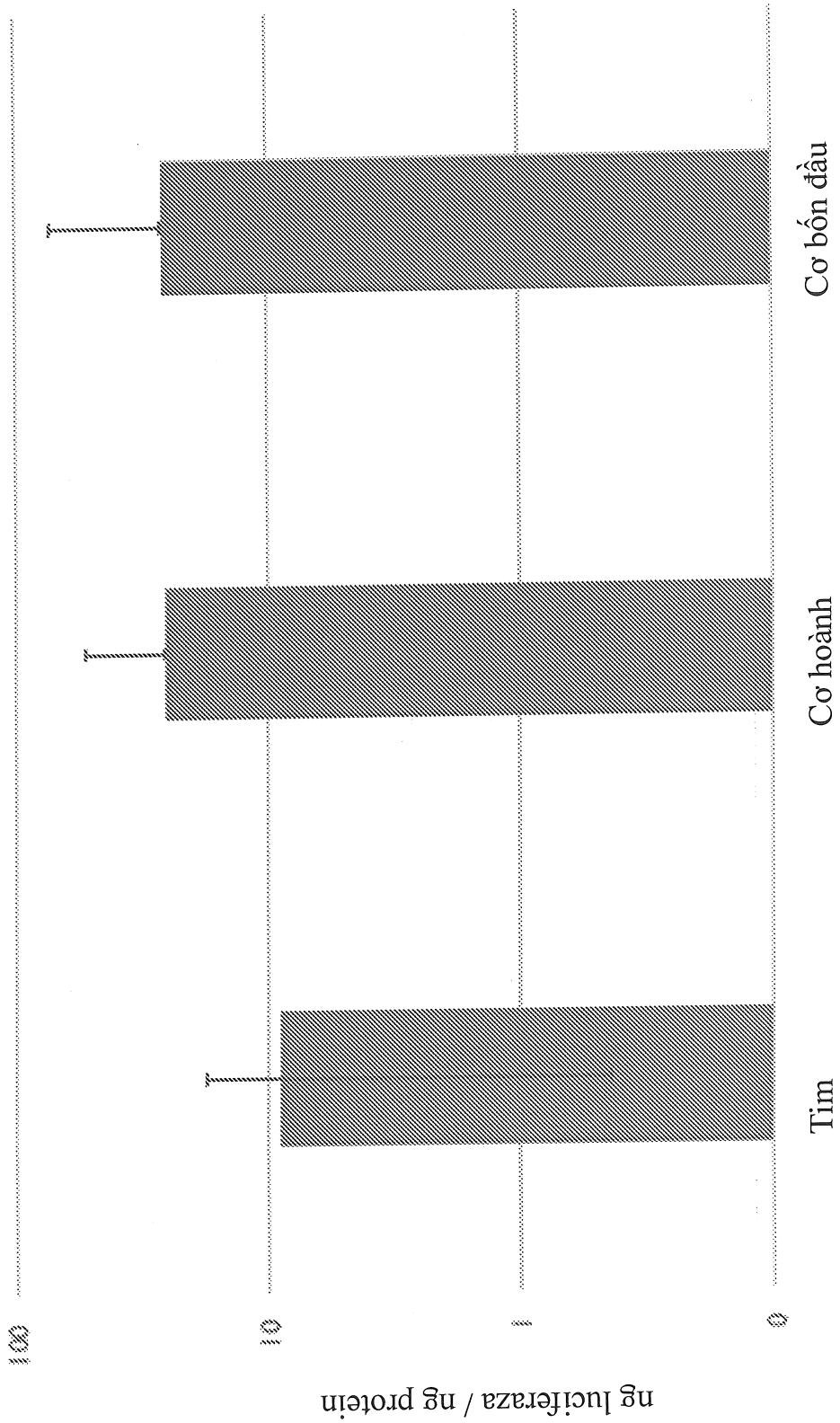


Fig. 12A

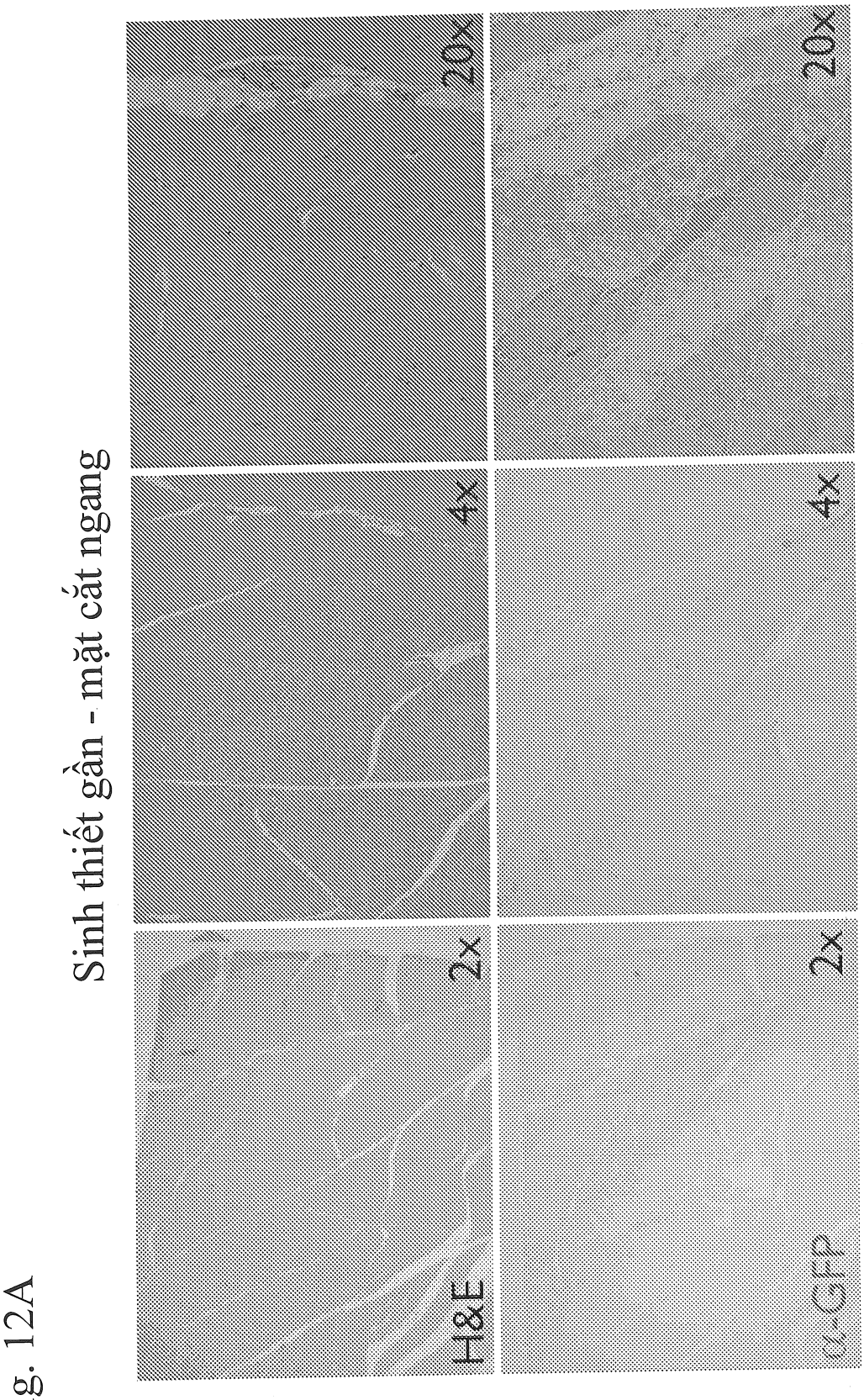
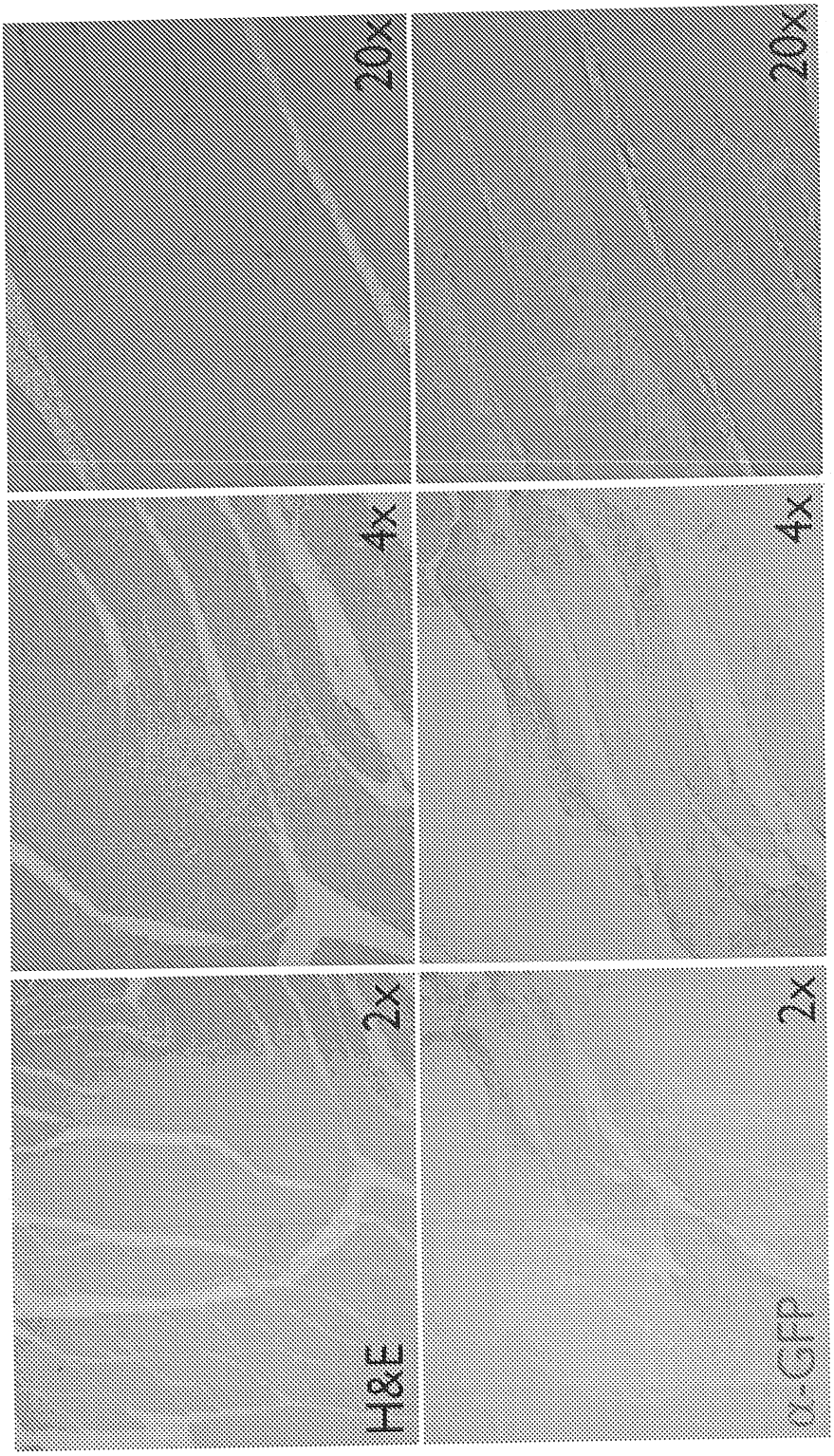


Fig. 12B

Sinh thiết xa - mặt cắt dọc



DANH SÁCH TRÌNH TỰ

<110> 4D Molecular Therapeutics Inc.

<120> Protein capsit virut adeno kết hợp biến thể và virion virut adeno kết hợp tái tổ hợp lây nhiễm

<130> 090400-5009 WO

<150> US 62/560,901

<151> 2017-09-20

<160> 69

<170> PatentIn version 3.5

<210> 1

<211> 736

<212> PRT

<213> Virut kết hợp adeno 1

<400> 1

```
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1               5                  10                  15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
            20                  25                  30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
        35                  40                  45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
    50                  55                  60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65                  70                  75                  80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
                85                  90                  95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
            100                 105                 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
        115                 120                 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
    130                 135                 140
```

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145 150 155 160

Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
180 185 190

Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His
260 265 270

Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
275 280 285

His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
290 295 300

Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln
305 310 315 320

Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
325 330 335

Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
340 345 350

Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
355 360 365

Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
370 375 380

Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
385 390 395 400

Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
405 410 415

Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
420 425 430

Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg
435 440 445

Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser
450 455 460

Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro
465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn
485 490 495

Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn
500 505 510

Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys
515 520 525

Asp Asp Glu Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly
530 535 540

Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile
545 550 555 560

Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg
565 570 575

Phe Gly Thr Val Ala Val Asn Phe Gln Ser Ser Ser Thr Asp Pro Ala

580 585 590

Thr Gly Asp Val His Ala Met Gly Ala Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625 630 635 640

Lys Asn Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Val Gln Tyr Thr Ser Asn
690 695 700

Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gly Leu
705 710 715 720

Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu
725 730 735

<210> 2
<211> 735
<212> PRT
<213> virut kết hợp adeno 2
<400> 2
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Arg Gln Ala Ala Thr
580 585 590

Ala Asp Val Asn Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp
595 600 605

Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr
610 615 620

Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys
625 630 635 640

His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn
645 650 655

Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln
660 665 670

Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
675 680 685

Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr

690 695 700

Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr
705 710 715 720

Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 3
<211> 736
<212> PRT
<213> Virut kết hợp adeno 3A
<400> 3
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro
20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Ile Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Gly
130 135 140

Ala Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly

145 150 155 160

Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
180 185 190

Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

Arg Gly Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu
325 330 335

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr
435 440 445

Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser
450 455 460

Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn
485 490 495

Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn
500 505 510

Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly
530 535 540

Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
545 550 555 560

Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
565 570 575

Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr
580 585 590

Thr Gly Thr Val Asn His Gln Gly Ala Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625 630 635 640

Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 4
<211> 736
<212> PRT
<213> Virut kết hợp adeno 3B
<400> 4
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Val Pro Gln Pro
20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Arg Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Ile Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly
145 150 155 160

Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
180 185 190

Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr

260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr
435 440 445

Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser
450 455 460

Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro
465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn
485 490 495

Asn Asn Ser Asn Phe Pro Trp Thr Ala Ala Ser Lys Tyr His Leu Asn
500 505 510

Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Asp Asp Glu Glu Lys Phe Phe Pro Met His Gly Asn Leu Ile Phe Gly
530 535 540

Lys Glu Gly Thr Thr Ala Ser Asn Ala Glu Leu Asp Asn Val Met Ile
545 550 555 560

Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
565 570 575

Tyr Gly Thr Val Ala Asn Asn Leu Gln Ser Ser Asn Thr Ala Pro Thr
580 585 590

Thr Arg Thr Val Asn Asp Gln Gly Ala Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625 630 635 640

Lys His Pro Pro Pro Gln Ile Met Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asn Pro Pro Thr Thr Phe Ser Pro Ala Lys Phe Ala Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 5
<211> 734
<212> PRT
<213> Virut kết hợp adeno 4
<400> 5

Met Thr Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu
1 5 10 15

Gly Val Arg Glu Trp Trp Ala Leu Gln Pro Gly Ala Pro Lys Pro Lys
20 25 30

Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly
35 40 45

Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val
50 55 60

Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln
65 70 75 80

Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp
85 90 95

Ala Glu Phe Gln Gln Arg Leu Gln Gly Asp Thr Ser Phe Gly Gly Asn
100 105 110

Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro Leu
115 120 125

Gly Leu Val Glu Gln Ala Gly Glu Thr Ala Pro Gly Lys Lys Arg Pro
130 135 140

Leu Ile Glu Ser Pro Gln Gln Pro Asp Ser Ser Thr Gly Ile Gly Lys
145 150 155 160

Lys Gly Lys Gln Pro Ala Lys Lys Lys Leu Val Phe Glu Asp Glu Thr
165 170 175

Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Thr Ser Gly Ala Met Ser
180 185 190

Asp Asp Ser Glu Met Arg Ala Ala Ala Gly Gly Ala Ala Val Glu Gly
195 200 205

Gly Gln Gly Ala Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys
210 215 220

Asp Ser Thr Trp Ser Glu Gly His Val Thr Thr Thr Ser Thr Arg Thr
225 230 235 240

Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Arg Leu Gly Glu
245 250 255

Ser Leu Gln Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr
260 265 270

Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln
275 280 285

Arg Leu Ile Asn Asn Asn Trp Gly Met Arg Pro Lys Ala Met Arg Val
290 295 300

Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu
305 310 315 320

Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp
325 330 335

Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser
340 345 350

Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr
355 360 365

Cys Gly Leu Val Thr Gly Asn Thr Ser Gln Gln Gln Thr Asp Arg Asn

370 375 380

Ala Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly
385 390 395 400

Asn Asn Phe Glu Ile Thr Tyr Ser Phe Glu Lys Val Pro Phe His Ser
405 410 415

Met Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Ile
420 425 430

Asp Gln Tyr Leu Trp Gly Leu Gln Ser Thr Thr Thr Gly Thr Thr Leu
435 440 445

Asn Ala Gly Thr Ala Thr Thr Asn Phe Thr Lys Leu Arg Pro Thr Asn
450 455 460

Phe Ser Asn Phe Lys Lys Asn Trp Leu Pro Gly Pro Ser Ile Lys Gln
465 470 475 480

Gln Gly Phe Ser Lys Thr Ala Asn Gln Asn Tyr Lys Ile Pro Ala Thr
485 490 495

Gly Ser Asp Ser Leu Ile Lys Tyr Glu Thr His Ser Thr Leu Asp Gly
500 505 510

Arg Trp Ser Ala Leu Thr Pro Gly Pro Pro Met Ala Thr Ala Gly Pro
515 520 525

Ala Asp Ser Lys Phe Ser Asn Ser Gln Leu Ile Phe Ala Gly Pro Lys
530 535 540

Gln Asn Gly Asn Thr Ala Thr Val Pro Gly Thr Leu Ile Phe Thr Ser
545 550 555 560

Glu Glu Glu Leu Ala Ala Thr Asn Ala Thr Asp Thr Asp Met Trp Gly
565 570 575

Asn Leu Pro Gly Gly Asp Gln Ser Asn Ser Asn Leu Pro Thr Val Asp
580 585 590

Arg Leu Thr Ala Leu Gly Ala Val Pro Gly Met Val Trp Gln Asn Arg
595 600 605

Asp Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp
610 615 620

Gly His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu Lys His
625 630 635 640

Pro Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala Asn Pro
645 650 655

Ala Thr Thr Phe Ser Ser Thr Pro Val Asn Ser Phe Ile Thr Gln Tyr
660 665 670

Ser Thr Gly Gln Val Ser Val Gln Ile Asp Trp Glu Ile Gln Lys Glu
675 680 685

Arg Ser Lys Arg Trp Asn Pro Glu Val Gln Phe Thr Ser Asn Tyr Gly
690 695 700

Gln Gln Asn Ser Leu Leu Trp Ala Pro Asp Ala Ala Gly Lys Tyr Thr
705 710 715 720

Glu Pro Arg Ala Ile Gly Thr Arg Tyr Leu Thr His His Leu
725 730

<210> 6
<211> 724
<212> PRT
<213> Virut kết hợp adeno 5
<400> 6
Met Ser Phe Val Asp His Pro Pro Asp Trp Leu Glu Glu Val Gly Glu
1 5 10 15

Gly Leu Arg Glu Phe Leu Gly Leu Glu Ala Gly Pro Pro Lys Pro Lys
20 25 30

Pro Asn Gln Gln His Gln Asp Gln Ala Arg Gly Leu Val Leu Pro Gly
35 40 45

Tyr Asn Tyr Leu Gly Pro Gly Asn Gly Leu Asp Arg Gly Glu Pro Val
50 55 60

Asn Arg Ala Asp Glu Val Ala Arg Glu His Asp Ile Ser Tyr Asn Glu
65 70 75 80

Gln Leu Glu Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp
85 90 95

Ala Glu Phe Gln Glu Lys Leu Ala Asp Asp Thr Ser Phe Gly Gly Asn
100 105 110

Leu Gly Lys Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro Phe
115 120 125

Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Thr Gly Lys Arg Ile
130 135 140

Asp Asp His Phe Pro Lys Arg Lys Lys Ala Arg Thr Glu Glu Asp Ser
145 150 155 160

Lys Pro Ser Thr Ser Ser Asp Ala Glu Ala Gly Pro Ser Gly Ser Gln
165 170 175

Gln Leu Gln Ile Pro Ala Gln Pro Ala Ser Ser Leu Gly Ala Asp Thr
180 185 190

Met Ser Ala Gly Gly Gly Gly Pro Leu Gly Asp Asn Asn Gln Gly Ala
195 200 205

Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys Asp Ser Thr Trp
210 215 220

Met Gly Asp Arg Val Val Thr Lys Ser Thr Arg Thr Trp Val Leu Pro
225 230 235 240

Ser Tyr Asn Asn His Gln Tyr Arg Glu Ile Lys Ser Gly Ser Val Asp
245 250 255

Gly Ser Asn Ala Asn Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr
260 265 270

Phe Asp Phe Asn Arg Phe His Ser His Trp Ser Pro Arg Asp Trp Gln
275 280 285

Arg Leu Ile Asn Asn Tyr Trp Gly Phe Arg Pro Arg Ser Leu Arg Val
290 295 300

Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Val Gln Asp Ser Thr
305 310 315 320

Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp
325 330 335

Asp Asp Tyr Gln Leu Pro Tyr Val Val Gly Asn Gly Thr Glu Gly Cys
340 345 350

Leu Pro Ala Phe Pro Pro Gln Val Phe Thr Leu Pro Gln Tyr Gly Tyr
355 360 365

Ala Thr Leu Asn Arg Asp Asn Thr Glu Asn Pro Thr Glu Arg Ser Ser
370 375 380

Phe Phe Cys Leu Glu Tyr Phe Pro Ser Lys Met Leu Arg Thr Gly Asn
385 390 395 400

Asn Phe Glu Phe Thr Tyr Asn Phe Glu Glu Val Pro Phe His Ser Ser
405 410 415

Phe Ala Pro Ser Gln Asn Leu Phe Lys Leu Ala Asn Pro Leu Val Asp
420 425 430

Gln Tyr Leu Tyr Arg Phe Val Ser Thr Asn Asn Thr Gly Gly Val Gln
435 440 445

Phe Asn Lys Asn Leu Ala Gly Arg Tyr Ala Asn Thr Tyr Lys Asn Trp
450 455 460

Phe Pro Gly Pro Met Gly Arg Thr Gln Gly Trp Asn Leu Gly Ser Gly
465 470 475 480

Val Asn Arg Ala Ser Val Ser Ala Phe Ala Thr Thr Asn Arg Met Glu

485 490 495

Leu Glu Gly Ala Ser Tyr Gln Val Pro Pro Gln Pro Asn Gly Met Thr
500 505 510

Asn Asn Leu Gln Gly Ser Asn Thr Tyr Ala Leu Glu Asn Thr Met Ile
515 520 525

Phe Asn Ser Gln Pro Ala Asn Pro Gly Thr Thr Ala Thr Tyr Leu Glu
530 535 540

Gly Asn Met Leu Ile Thr Ser Glu Ser Glu Thr Gln Pro Val Asn Arg
545 550 555 560

Val Ala Tyr Asn Val Gly Gly Gln Met Ala Thr Asn Asn Gln Ser Ser
565 570 575

Thr Thr Ala Pro Ala Thr Gly Thr Tyr Asn Leu Gln Glu Ile Val Pro
580 585 590

Gly Ser Val Trp Met Glu Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp
595 600 605

Ala Lys Ile Pro Glu Thr Gly Ala His Phe His Pro Ser Pro Ala Met
610 615 620

Gly Gly Phe Gly Leu Lys His Pro Pro Pro Met Met Leu Ile Lys Asn
625 630 635 640

Thr Pro Val Pro Gly Asn Ile Thr Ser Phe Ser Asp Val Pro Val Ser
645 650 655

Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Thr Val Glu Met Glu
660 665 670

Trp Glu Leu Lys Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln
675 680 685

Tyr Thr Asn Asn Tyr Asn Asp Pro Gln Phe Val Asp Phe Ala Pro Asp
690 695 700

Ser Thr Gly Glu Tyr Arg Thr Thr Arg Pro Ile Gly Thr Arg Tyr Leu
705 710 715 720

Thr Arg Pro Leu

<210> 7
<211> 736
<212> PRT
<213> Virut kết hợp adeno 6
<400> 7
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Phe Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly
145 150 155 160

Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro
180 185 190

Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His
260 265 270

Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
275 280 285

His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
290 295 300

Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln
305 310 315 320

Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn
325 330 335

Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro
340 345 350

Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
355 360 365

Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
370 375 380

Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
385 390 395 400

Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe
405 410 415

Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
420 425 430

Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg
435 440 445

Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser
450 455 460

Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro
465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn
485 490 495

Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn
500 505 510

Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys
515 520 525

Asp Asp Lys Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly
530 535 540

Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile
545 550 555 560

Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg
565 570 575

Phe Gly Thr Val Ala Val Asn Leu Gln Ser Ser Ser Thr Asp Pro Ala
580 585 590

Thr Gly Asp Val His Val Met Gly Ala Leu Pro Gly Met Val Trp Gln

595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Val Gln Tyr Thr Ser Asn
690 695 700

Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gly Leu
705 710 715 720

Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu
725 730 735

<210> 8
<211> 737
<212> PRT
<213> Virut kết hợp adeno 7
<400> 8
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asn Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Ala Lys Lys Arg
130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
165 170 175

Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro
180 185 190

Pro Ala Ala Pro Ser Ser Val Gly Ser Gly Thr Val Ala Ala Gly Gly
195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn
210 215 220

Ala Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240

Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255

Leu Tyr Lys Gln Ile Ser Ser Glu Thr Ala Gly Ser Thr Asn Asp Asn
260 265 270

Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Lys Leu Arg Phe Lys Leu Phe Asn Ile
305 310 315 320

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

Asn Leu Thr Ser Thr Ile Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn
370 375 380

Gly Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr Ser
405 410 415

Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala
435 440 445

Arg Thr Gln Ser Asn Pro Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln
450 455 460

Phe Tyr Gln Gly Gly Pro Ser Thr Met Ala Glu Gln Ala Lys Asn Trp
465 470 475 480

Leu Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp
485 490 495

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510

Leu Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525

His Lys Asp Asp Glu Asp Arg Phe Phe Pro Ser Ser Gly Val Leu Ile
530 535 540

Phe Gly Lys Thr Gly Ala Thr Asn Lys Thr Thr Leu Glu Asn Val Leu
545 550 555 560

Met Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu
565 570 575

Glu Tyr Gly Ile Val Ser Ser Asn Leu Gln Ala Ala Asn Thr Ala Ala
580 585 590

Gln Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp
595 600 605

Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro
610 615 620

His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly
625 630 635 640

Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro
645 650 655

Ala Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile
660 665 670

Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu
675 680 685

Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser
690 695 700

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

705 710 715 720

Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn
725 730 735

Leu

<210> 9
<211> 738
<212> PRT
<213> Virut kết hợp adeno 8
<400> 9
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Gln Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile

145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
165 170 175

Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro
180 185 190

Pro Ala Ala Pro Ser Gly Val Gly Pro Asn Thr Met Ala Ala Gly Gly
195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220

Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240

Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255

Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ala Thr Asn Asp
260 265 270

Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285

Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300

Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn
305 310 315 320

Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335

Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350

Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365

Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380

Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400

Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Thr Tyr
405 410 415

Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430

Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445

Ser Arg Thr Gln Thr Thr Gly Gly Thr Ala Asn Thr Gln Thr Leu Gly
450 455 460

Phe Ser Gln Gly Gly Pro Asn Thr Met Ala Asn Gln Ala Lys Asn Trp
465 470 475 480

Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Gly
485 490 495

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Ala Gly Thr Lys Tyr His
500 505 510

Leu Asn Gly Arg Asn Ser Leu Ala Asn Pro Gly Ile Ala Met Ala Thr
515 520 525

His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Asn Gly Ile Leu Ile
530 535 540

Phe Gly Lys Gln Asn Ala Ala Arg Asp Asn Ala Asp Tyr Ser Asp Val
545 550 555 560

Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575

Glu Glu Tyr Gly Ile Val Ala Asp Asn Leu Gln Gln Gln Asn Thr Ala
580 585 590

Pro Gln Ile Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655

Pro Ala Asp Pro Pro Thr Thr Phe Asn Gln Ser Lys Leu Asn Ser Phe
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Ser Val Asp Phe Ala Val Asn Thr Glu
705 710 715 720

Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735

Asn Leu

<210> 10
<211> 736
<212> PRT
<213> Virut kết hợp adeno 9
<400> 10
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

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

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 11
<211> 738
<212> PRT
<213> Virut kết hợp adeno 10
<400> 11
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
165 170 175

Thr Gly Glu Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190

Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220

Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240

Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255

Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270

Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285

Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300

Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn
305 310 315 320

Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335

Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350

Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe

355 360 365

Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380

Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400

Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
405 410 415

Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430

Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445

Ser Arg Thr Gln Ser Thr Gly Gly Thr Gln Gly Thr Gln Gln Leu Leu
450 455 460

Phe Ser Gln Ala Gly Pro Ala Asn Met Ser Ala Gln Ala Lys Asn Trp
465 470 475 480

Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
485 490 495

Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510

Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525

His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
530 535 540

Phe Gly Lys Gln Gly Ala Gly Arg Asp Asn Val Asp Tyr Ser Ser Val
545 550 555 560

Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575

Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Ala Asn Thr Gly
580 585 590

Pro Ile Val Gly Asn Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655

Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu
705 710 715 720

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735

Asn Leu

<210> 12
<211> 738
<212> PRT
<213> Virut kết hợp adeno rh10
<400> 12
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
165 170 175

Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190

Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220

Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Gly Val
225 230 235 240

Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255

Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270

Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285

Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300

Asn Asn Trp Gly Phe Arg Pro Lys Ser Leu Asn Phe Lys Leu Phe Asn
305 310 315 320

Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335

Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350

Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365

Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380

Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
385 390 395 400

Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
405 410 415

Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
420 425 430

Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
435 440 445

Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu

450 455 460

Phe Ser Gln Ala Gly Pro Asn Asn Met Ser Ala Gln Ala Lys Asn Trp
465 470 475 480

Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
485 490 495

Gln Asn Asp Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
500 505 510

Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
515 520 525

His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Met
530 535 540

Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Ser Val
545 550 555 560

Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
565 570 575

Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Gln Gln Asn Ala Ala
580 585 590

Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655

Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Asp
705 710 715 720

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735

Asn Leu

<210> 13
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 13
Asn Lys Ile Gln Arg Thr Asp
1 5

<210> 14
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 14
Asn Lys Thr Thr Asn Lys Asp
1 5

<210> 15
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại

<400> 15
Thr Asn Lys Ile Gly Val Thr
1 5

<210> 16
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 16
Gly Asn Leu Thr Lys Gly Asn
1 5

<210> 17
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 17
Asn Thr Val Lys Leu Ser Thr
1 5

<210> 18
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 18
Ser Asn Thr Val Lys Ala Ile
1 5

<210> 19
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại

<400> 19
Ala Ser Asn Ile Thr Lys Ala
1 5

<210> 20
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 20
Asp Asn Thr Val Thr Arg Ser
1 5

<210> 21
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 21
Asn Lys Ile Ser Ala Lys Asp
1 5

<210> 22
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 22
Asn Gln Asp Tyr Thr Lys Thr
1 5

<210> 23
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại

<400> 23
Gln Ala Asp Thr Thr Lys Asn
1 5

<210> 24
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 24
Thr Asn Arg Thr Ser Pro Asp
1 5

<210> 25
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 25
Ser Asn Thr Thr Gln Lys Thr
1 5

<210> 26
<211> 7
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 26
Ala Ser Asp Ser Thr Lys Ala
1 5

<210> 27
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại

<400> 27
Leu Ala Asn Lys Ile Gln Arg Thr Asp Ala
1 5 10

<210> 28
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 28
Leu Ala Asn Lys Thr Thr Asn Lys Asp Ala
1 5 10

<210> 29
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 29
Leu Ala Thr Asn Lys Ile Gly Val Thr Ala
1 5 10

<210> 30
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 30
Leu Ala Gly Asn Leu Thr Lys Gly Asn Ala
1 5 10

<210> 31
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại

<400> 31
Leu Ala Asn Thr Val Lys Leu Ser Thr Ala
1 5 10

<210> 32
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 32
Leu Ala Ser Asn Thr Val Lys Ala Ile Ala
1 5 10

<210> 33
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 33
Leu Ala Ala Ser Asn Ile Thr Lys Ala Ala
1 5 10

<210> 34
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 34
Leu Ala Asp Asn Thr Val Thr Arg Ser Ala
1 5 10

<210> 35
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại

<400> 35
Leu Ala Asn Lys Ile Ser Ala Lys Asp Ala
1 5 10

<210> 36
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 36
Leu Ala Asn Gln Asp Tyr Thr Lys Thr Ala
1 5 10

<210> 37
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 37
Leu Ala Thr Asn Lys Ile Gly Val Thr Ser
1 5 10

<210> 38
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 38
Leu Ala Thr Asn Lys Ile Gly Val Thr Ala
1 5 10

<210> 39
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại

<400> 39
Leu Ala Gln Ala Asp Thr Thr Lys Asn Ala
1 5 10

<210> 40
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 40
Leu Ala Thr Asn Arg Thr Ser Pro Asp Ala
1 5 10

<210> 41
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 41
Leu Ala Ser Asn Thr Thr Gln Lys Thr Ala
1 5 10

<210> 42
<211> 10
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Chèn thêm peptit khác loại
<400> 42
Leu Ala Ala Ser Asp Ser Thr Lys Ala Ala
1 5 10

<210> 43
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể

<400> 43

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser

210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr

435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asn Lys Ile
580 585 590

Gln Arg Thr Asp Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Ile Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 44
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 44
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Lys Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asp Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Ser Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asn Lys Ile
580 585 590

Gln Arg Thr Asp Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Val Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Gln

740 745

<210> 45
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 45

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser

385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asn Lys Ile
580 585 590

Gln Arg Thr Asp Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 46
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 46
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Thr Asn Lys
580 585 590

Ile Gly Val Thr Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp

690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Ile Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 47
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 47

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr

340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Thr Asn Lys
580 585 590

Ile Gly Val Thr Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 48
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>

<223> Capsit AAV biến thể

<400> 48

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asn Lys Thr
580 585 590

Thr Asn Lys Asp Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile

645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Ile Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 49
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 49

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp

290 295 300

Gly Phe Arg Pro Lys Arg Leu Lys Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asp Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Ser Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asn Lys Thr
580 585 590

Thr Asn Lys Asp Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Val Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Gln
740 745

<210> 50
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 50

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asn Lys Thr
580 585 590

Thr Asn Lys Asp Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln

595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 51
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 51

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu

245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Gly Asn Leu
580 585 590

Thr Lys Gly Asn Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 52
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 52
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr

545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asn Thr Val
580 585 590

Lys Leu Ser Thr Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 53
<211> 745
<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Capsit AAV biến thể

<400> 53

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly

195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Ser Asn Thr
580 585 590

Val Lys Ala Ile Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 54
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 54
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly

500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Ala Ser Asn
580 585 590

Ile Thr Lys Ala Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 55
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 55
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly

145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asp Asn Thr
580 585 590

Val Thr Arg Ser Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 56
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 56
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

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

450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asn Lys Ile
580 585 590

Ser Ala Lys Asp Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 57
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 57
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Asn Gln Asp
580 585 590

Tyr Thr Lys Thr Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 58

<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 58

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu

405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Gln Ala Asp
580 585 590

Thr Thr Lys Asn Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 59
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 59
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Thr Asn Arg
580 585 590

Thr Ser Pro Asp Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Ile Asn Val

705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 60
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 60

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp

355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Ser Asn Thr
580 585 590

Thr Gln Lys Thr Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser
660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 61
<211> 745
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể
<400> 61
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser

1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

Ala Ala Pro Ser Gly Leu Gly Thr Asn Thr Met Ala Thr Gly Ser Gly
195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His
275 280 285

Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp
290 295 300

Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val
305 310 315 320

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser
385 390 395 400

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr
435 440 445

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly
465 470 475 480

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly
500 505 510

Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp
515 520 525

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Leu Ala Ala Ser Asp
580 585 590

Ser Thr Lys Ala Ala Arg Gln Ala Ala Thr Ala Asp Val Asn Thr Gln
595 600 605

Gly Val Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln
610 615 620

Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro
625 630 635 640

Ser Pro Leu Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Gln Ile
645 650 655

Leu Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Thr Thr Phe Ser

660 665 670

Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val
675 680 685

Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp
690 695 700

Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Asn Lys Ser Val Asn Val
705 710 715 720

Asp Phe Thr Val Asp Thr Asn Gly Val Tyr Ser Glu Pro Arg Pro Ile
725 730 735

Gly Thr Arg Tyr Leu Thr Arg Asn Leu
740 745

<210> 62
<211> 725
<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể chimera
<400> 62
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Phe Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Thr Gly Lys Arg
130 135 140

Ile Asp Asp His Phe Pro Lys Arg Lys Lys Ala Arg Thr Glu Glu Asp
145 150 155 160

Ser Lys Pro Ser Thr Ser Ser Asp Ala Glu Ala Gly Pro Ser Gly Ser
165 170 175

Gln Gln Leu Gln Ile Pro Ala Gln Pro Ala Ser Ser Leu Gly Ala Asp
180 185 190

Thr Met Ser Ala Gly Gly Gly Gly Pro Leu Gly Asp Asn Asn Gln Gly
195 200 205

Ala Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys Asp Ser Thr
210 215 220

Trp Met Gly Asp Arg Ile Val Thr Lys Ser Thr Arg Thr Trp Val Leu
225 230 235 240

Pro Ser Tyr Asn Asn His Gln Tyr Arg Glu Ile Lys Ser Gly Ser Val
245 250 255

Asp Gly Ser Asn Ala Asn Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly
260 265 270

Tyr Phe Asp Phe Asn Arg Phe His Ser His Trp Ser Pro Arg Asp Trp
275 280 285

Gln Arg Leu Ile Asn Asn Tyr Trp Gly Phe Arg Pro Arg Ser Leu Arg
290 295 300

Val Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Val Gln Asp Ser

305 310 315 320

Thr Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Val Gln Val Phe Thr
325 330 335

Asp Asp Asp Tyr Gln Leu Pro Tyr Val Val Gly Asn Gly Thr Glu Gly
340 345 350

Cys Leu Pro Ala Phe Pro Pro Gln Val Phe Thr Leu Pro Gln Tyr Gly
355 360 365

Tyr Ala Thr Leu Asn Arg Asp Asn Thr Glu Asn Pro Thr Glu Arg Ser
370 375 380

Ser Phe Phe Cys Leu Glu Tyr Phe Pro Ser Lys Met Leu Arg Thr Gly
385 390 395 400

Asn Asn Phe Glu Phe Thr Tyr Asn Phe Glu Glu Val Pro Phe His Ser
405 410 415

Ser Phe Ala Pro Ser Gln Asn Leu Phe Lys Leu Ala Asn Pro Leu Val
420 425 430

Asp Gln Tyr Leu Tyr Arg Phe Val Ser Thr Asn Asn Thr Gly Gly Val
435 440 445

Gln Phe Asn Lys Asn Leu Ala Gly Arg Tyr Ala Asn Thr Tyr Lys Asn
450 455 460

Trp Phe Pro Gly Pro Met Gly Arg Thr Gln Gly Trp Asn Leu Gly Ser
465 470 475 480

Gly Val Asn Arg Ala Ser Val Ser Ala Phe Thr Thr Thr Asn Arg Met
485 490 495

Glu Leu Glu Gly Ala Ser Tyr Gln Val Pro Pro Gln Pro Asn Gly Met
500 505 510

Thr Asn Asn Leu Gln Gly Ser Asn Thr Tyr Ala Leu Glu Asn Thr Met
515 520 525

Ile Phe Asn Ser Gln Pro Ala Asn Pro Gly Thr Thr Ala Thr Tyr Leu
530 535 540

Glu Gly Asn Met Leu Ile Thr Ser Glu Ser Glu Thr Gln Pro Val Asn
545 550 555 560

Arg Val Ala Tyr Asn Val Gly Gly Gln Met Ala Thr Asn Asn Gln Ser
565 570 575

Ser Thr Thr Ala Pro Thr Thr Gly Thr Tyr Asn Leu Gln Glu Ile Val
580 585 590

Pro Gly Ser Val Trp Met Glu Arg Asp Val Tyr Leu Gln Gly Pro Ile
595 600 605

Trp Ala Lys Ile Pro Glu Thr Gly Ala His Phe His Pro Ser Pro Ala
610 615 620

Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Met Met Leu Ile Lys
625 630 635 640

Asn Thr Pro Val Pro Gly Asn Ile Thr Ser Phe Ser Asp Val Pro Val
645 650 655

Ser Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Thr Val Glu Met
660 665 670

Glu Trp Glu Leu Lys Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile
675 680 685

Gln Tyr Thr Asn Asn Tyr Asn Asp Pro Gln Phe Val Asp Phe Ala Pro
690 695 700

Asp Ser Thr Gly Glu Tyr Arg Thr Thr Arg Pro Ile Gly Thr Arg Tyr
705 710 715 720

Leu Thr Arg Pro Leu
725

<210> 63
<211> 725

<212> PRT
<213> Trình tự nhân tạo
<220>
<223> Capsit AAV biến thể chimera
<400> 63

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser
1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro
20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Phe Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Thr Gly Lys Arg
130 135 140

Ile Asp Asp His Phe Pro Lys Arg Lys Lys Ala Arg Thr Glu Glu Asp
145 150 155 160

Ser Lys Pro Ser Thr Ser Ser Asp Ala Glu Ala Gly Pro Ser Gly Ser
165 170 175

Gln Gln Leu Gln Ile Pro Ala Gln Pro Ala Ser Ser Leu Gly Ala Asp
180 185 190

Thr Met Ser Ala Gly Gly Gly Gly Pro Leu Gly Asp Asn Asn Gln Gly
195 200 205

Ala Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys Asp Ser Thr
210 215 220

Trp Met Gly Asp Arg Ile Val Thr Lys Ser Thr Arg Thr Trp Val Leu
225 230 235 240

Pro Ser Tyr Asn Asn His Gln Tyr Arg Glu Ile Lys Ser Gly Ser Val
245 250 255

Asp Gly Ser Asn Ala Asn Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly
260 265 270

Tyr Phe Asp Phe Asn Arg Phe His Ser His Trp Ser Pro Arg Asp Trp
275 280 285

Gln Arg Leu Ile Asn Asn Tyr Trp Gly Phe Arg Pro Arg Ser Leu Arg
290 295 300

Val Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Val Gln Asp Ser
305 310 315 320

Thr Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Val Gln Val Phe Thr
325 330 335

Asp Asp Asp Tyr Gln Leu Pro Tyr Val Val Gly Asn Gly Thr Glu Gly
340 345 350

Cys Leu Pro Ala Phe Pro Pro Gln Val Phe Thr Leu Pro Gln Tyr Gly
355 360 365

Tyr Ala Thr Leu Asn Arg Asp Asn Thr Glu Asn Pro Thr Glu Arg Ser
370 375 380

Ser Phe Phe Cys Leu Glu Tyr Phe Pro Ser Lys Met Leu Arg Thr Gly
385 390 395 400

Asn Asn Phe Glu Phe Thr Tyr Asn Phe Glu Glu Val Pro Phe His Ser
405 410 415

Ser Phe Ala Pro Ser Gln Asn Leu Phe Lys Leu Ala Asn Pro Leu Val
420 425 430

Asp Gln Tyr Leu Tyr Arg Phe Val Ser Thr Asn Asn Thr Gly Gly Val
435 440 445

Gln Phe Asn Lys Asn Leu Ala Gly Arg Tyr Ala Asn Thr Tyr Lys Asn
450 455 460

Trp Phe Pro Gly Pro Met Gly Arg Thr Gln Gly Trp Asn Leu Gly Ser
465 470 475 480

Gly Val Asn Arg Ala Ser Val Ser Ala Phe Thr Thr Thr Asn Arg Met
485 490 495

Glu Leu Glu Gly Ala Ser Tyr Gln Val Pro Pro Gln Pro Asn Gly Met
500 505 510

Thr Asn Asn Leu Gln Gly Ser Asn Thr Tyr Ala Leu Glu Asn Thr Met
515 520 525

Ile Phe Asn Ser Gln Pro Ala Asn Pro Gly Thr Thr Ala Thr Tyr Leu
530 535 540

Glu Gly Asn Met Leu Ile Thr Ser Glu Ser Glu Thr Gln Pro Val Asn
545 550 555 560

Arg Val Ala Tyr Asn Val Gly Gly Gln Met Ala Thr Asn Asn Gln Ser
565 570 575

Ser Thr Thr Ala Pro Thr Thr Gly Thr Tyr Asn Leu Gln Glu Ile Val
580 585 590

Pro Gly Ser Val Trp Met Glu Arg Asp Val Tyr Leu Gln Gly Pro Ile
595 600 605

Trp Ala Lys Ile Pro Glu Thr Gly Ala His Phe His Pro Ser Pro Ala
610 615 620

Met Gly Gly Phe Gly Leu Lys His Pro Pro Pro Met Met Leu Ile Lys

625 630 635 640

Asn Thr Pro Val Pro Gly Asn Ile Thr Ser Phe Ser Asp Val Pro Val
645 650 655

Ser Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Thr Val Glu Met
660 665 670

Glu Trp Glu Leu Lys Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile
675 680 685

Gln Tyr Thr Asn Asn Tyr Asn Asp Pro Gln Phe Val Asp Phe Ala Pro
690 695 700

Asp Ser Thr Gly Glu Tyr Arg Thr Thr Arg Pro Ile Gly Thr Arg Tyr
705 710 715 720

Leu Thr Arg Pro Leu
725

<210> 64
<211> 3685
<212> PRT
<213> Loài người
<400> 64

Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
20 25 30

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
50 55 60

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
65 70 75 80

Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser

85 90 95

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
145 150 155 160

Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His
165 170 175

Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala
180 185 190

Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly
195 200 205

Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp
210 215 220

Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro
225 230 235 240

Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg
245 250 255

Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met
260 265 270

His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg
275 280 285

Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala
290 295 300

Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln
305 310 315 320

His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu
325 330 335

Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu
340 345 350

Ser Trp Leu Leu Ser Ala Glu Asp Thr Leu Gln Ala Gln Gly Glu Ile
355 360 365

Ser Asn Asp Val Glu Val Val Lys Asp Gln Phe His Thr His Glu Gly
370 375 380

Tyr Met Met Asp Leu Thr Ala His Gln Gly Arg Val Gly Asn Ile Leu
385 390 395 400

Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu
405 410 415

Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Glu
420 425 430

Cys Leu Arg Val Ala Ser Met Glu Lys Gln Ser Asn Leu His Arg Val
435 440 445

Leu Met Asp Leu Gln Asn Gln Lys Leu Lys Glu Leu Asn Asp Trp Leu
450 455 460

Thr Lys Thr Glu Glu Arg Thr Arg Lys Met Glu Glu Glu Pro Leu Gly
465 470 475 480

Pro Asp Leu Glu Asp Leu Lys Arg Gln Val Gln Gln His Lys Val Leu
485 490 495

Gln Glu Asp Leu Glu Gln Glu Gln Val Arg Val Asn Ser Leu Thr His
500 505 510

Met Val Val Val Val Asp Glu Ser Ser Gly Asp His Ala Thr Ala Ala
515 520 525

Leu Glu Glu Gln Leu Lys Val Leu Gly Asp Arg Trp Ala Asn Ile Cys
530 535 540

Arg Trp Thr Glu Asp Arg Trp Val Leu Leu Gln Asp Ile Leu Leu Lys
545 550 555 560

Trp Gln Arg Leu Thr Glu Glu Gln Cys Leu Phe Ser Ala Trp Leu Ser
565 570 575

Glu Lys Glu Asp Ala Val Asn Lys Ile His Thr Thr Gly Phe Lys Asp
580 585 590

Gln Asn Glu Met Leu Ser Ser Leu Gln Lys Leu Ala Val Leu Lys Ala
595 600 605

Asp Leu Glu Lys Lys Lys Gln Ser Met Gly Lys Leu Tyr Ser Leu Lys
610 615 620

Gln Asp Leu Leu Ser Thr Leu Lys Asn Lys Ser Val Thr Gln Lys Thr
625 630 635 640

Glu Ala Trp Leu Asp Asn Phe Ala Arg Cys Trp Asp Asn Leu Val Gln
645 650 655

Lys Leu Glu Lys Ser Thr Ala Gln Ile Ser Gln Ala Val Thr Thr Thr
660 665 670

Gln Pro Ser Leu Thr Gln Thr Thr Val Met Glu Thr Val Thr Thr Val
675 680 685

Thr Thr Arg Glu Gln Ile Leu Val Lys His Ala Gln Glu Glu Leu Pro
690 695 700

Pro Pro Pro Pro Gln Lys Lys Arg Gln Ile Thr Val Asp Ser Glu Ile
705 710 715 720

Arg Lys Arg Leu Asp Val Asp Ile Thr Glu Leu His Ser Trp Ile Thr
725 730 735

Arg Ser Glu Ala Val Leu Gln Ser Pro Glu Phe Ala Ile Phe Arg Lys

740 745 750

Glu Gly Asn Phe Ser Asp Leu Lys Glu Lys Val Asn Ala Ile Glu Arg
755 760 765

Glu Lys Ala Glu Lys Phe Arg Lys Leu Gln Asp Ala Ser Arg Ser Ala
770 775 780

Gln Ala Leu Val Glu Gln Met Val Asn Glu Gly Val Asn Ala Asp Ser
785 790 795 800

Ile Lys Gln Ala Ser Glu Gln Leu Asn Ser Arg Trp Ile Glu Phe Cys
805 810 815

Gln Leu Leu Ser Glu Arg Leu Asn Trp Leu Glu Tyr Gln Asn Asn Ile
820 825 830

Ile Ala Phe Tyr Asn Gln Leu Gln Gln Leu Glu Gln Met Thr Thr Thr
835 840 845

Ala Glu Asn Trp Leu Lys Ile Gln Pro Thr Thr Pro Ser Glu Pro Thr
850 855 860

Ala Ile Lys Ser Gln Leu Lys Ile Cys Lys Asp Glu Val Asn Arg Leu
865 870 875 880

Ser Gly Leu Gln Pro Gln Ile Glu Arg Leu Lys Ile Gln Ser Ile Ala
885 890 895

Leu Lys Glu Lys Gly Gln Gly Pro Met Phe Leu Asp Ala Asp Phe Val
900 905 910

Ala Phe Thr Asn His Phe Lys Gln Val Phe Ser Asp Val Gln Ala Arg
915 920 925

Glu Lys Glu Leu Gln Thr Ile Phe Asp Thr Leu Pro Pro Met Arg Tyr
930 935 940

Gln Glu Thr Met Ser Ala Ile Arg Thr Trp Val Gln Gln Ser Glu Thr
945 950 955 960

Lys Leu Ser Ile Pro Gln Leu Ser Val Thr Asp Tyr Glu Ile Met Glu
965 970 975

Gln Arg Leu Gly Glu Leu Gln Ala Leu Gln Ser Ser Leu Gln Glu Gln
980 985 990

Gln Ser Gly Leu Tyr Tyr Leu Ser Thr Thr Val Lys Glu Met Ser Lys
995 1000 1005

Lys Ala Pro Ser Glu Ile Ser Arg Lys Tyr Gln Ser Glu Phe Glu
1010 1015 1020

Glu Ile Glu Gly Arg Trp Lys Lys Leu Ser Ser Gln Leu Val Glu
1025 1030 1035

His Cys Gln Lys Leu Glu Glu Gln Met Asn Lys Leu Arg Lys Ile
1040 1045 1050

Gln Asn His Ile Gln Thr Leu Lys Lys Trp Met Ala Glu Val Asp
1055 1060 1065

Val Phe Leu Lys Glu Glu Trp Pro Ala Leu Gly Asp Ser Glu Ile
1070 1075 1080

Leu Lys Lys Gln Leu Lys Gln Cys Arg Leu Leu Val Ser Asp Ile
1085 1090 1095

Gln Thr Ile Gln Pro Ser Leu Asn Ser Val Asn Glu Gly Gly Gln
1100 1105 1110

Lys Ile Lys Asn Glu Ala Glu Pro Glu Phe Ala Ser Arg Leu Glu
1115 1120 1125

Thr Glu Leu Lys Glu Leu Asn Thr Gln Trp Asp His Met Cys Gln
1130 1135 1140

Gln Val Tyr Ala Arg Lys Glu Ala Leu Lys Gly Gly Leu Glu Lys
1145 1150 1155

Thr Val Ser Leu Gln Lys Asp Leu Ser Glu Met His Glu Trp Met
1160 1165 1170

Thr Gln Ala Glu Glu Glu Tyr Leu Glu Arg Asp Phe Glu Tyr Lys
1175 1180 1185

Thr Pro Asp Glu Leu Gln Lys Ala Val Glu Glu Met Lys Arg Ala
1190 1195 1200

Lys Glu Glu Ala Gln Gln Lys Glu Ala Lys Val Lys Leu Leu Thr
1205 1210 1215

Glu Ser Val Asn Ser Val Ile Ala Gln Ala Pro Pro Val Ala Gln
1220 1225 1230

Glu Ala Leu Lys Lys Glu Leu Glu Thr Leu Thr Thr Asn Tyr Gln
1235 1240 1245

Trp Leu Cys Thr Arg Leu Asn Gly Lys Cys Lys Thr Leu Glu Glu
1250 1255 1260

Val Trp Ala Cys Trp His Glu Leu Leu Ser Tyr Leu Glu Lys Ala
1265 1270 1275

Asn Lys Trp Leu Asn Glu Val Glu Phe Lys Leu Lys Thr Thr Glu
1280 1285 1290

Asn Ile Pro Gly Gly Ala Glu Glu Ile Ser Glu Val Leu Asp Ser
1295 1300 1305

Leu Glu Asn Leu Met Arg His Ser Glu Asp Asn Pro Asn Gln Ile
1310 1315 1320

Arg Ile Leu Ala Gln Thr Leu Thr Asp Gly Gly Val Met Asp Glu
1325 1330 1335

Leu Ile Asn Glu Glu Leu Glu Thr Phe Asn Ser Arg Trp Arg Glu
1340 1345 1350

Leu His Glu Glu Ala Val Arg Arg Gln Lys Leu Leu Glu Gln Ser
1355 1360 1365

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

1370 1375 1380

Glu Ser Leu Thr Phe Ile Asp Lys Gln Leu Ala Ala Tyr Ile Ala
1385 1390 1395

Asp Lys Val Asp Ala Ala Gln Met Pro Gln Glu Ala Gln Lys Ile
1400 1405 1410

Gln Ser Asp Leu Thr Ser His Glu Ile Ser Leu Glu Glu Met Lys
1415 1420 1425

Lys His Asn Gln Gly Lys Glu Ala Ala Gln Arg Val Leu Ser Gln
1430 1435 1440

Ile Asp Val Ala Gln Lys Lys Leu Gln Asp Val Ser Met Lys Phe
1445 1450 1455

Arg Leu Phe Gln Lys Pro Ala Asn Phe Glu Gln Arg Leu Gln Glu
1460 1465 1470

Ser Lys Met Ile Leu Asp Glu Val Lys Met His Leu Pro Ala Leu
1475 1480 1485

Glu Thr Lys Ser Val Glu Gln Glu Val Val Gln Ser Gln Leu Asn
1490 1495 1500

His Cys Val Asn Leu Tyr Lys Ser Leu Ser Glu Val Lys Ser Glu
1505 1510 1515

Val Glu Met Val Ile Lys Thr Gly Arg Gln Ile Val Gln Lys Lys
1520 1525 1530

Gln Thr Glu Asn Pro Lys Glu Leu Asp Glu Arg Val Thr Ala Leu
1535 1540 1545

Lys Leu His Tyr Asn Glu Leu Gly Ala Lys Val Thr Glu Arg Lys
1550 1555 1560

Gln Gln Leu Glu Lys Cys Leu Lys Leu Ser Arg Lys Met Arg Lys
1565 1570 1575

Glu Met Asn Val Leu Thr Glu Trp Leu Ala Ala Thr Asp Met Glu
1580 1585 1590

Leu Thr Lys Arg Ser Ala Val Glu Gly Met Pro Ser Asn Leu Asp
1595 1600 1605

Ser Glu Val Ala Trp Gly Lys Ala Thr Gln Lys Glu Ile Glu Lys
1610 1615 1620

Gln Lys Val His Leu Lys Ser Ile Thr Glu Val Gly Glu Ala Leu
1625 1630 1635

Lys Thr Val Leu Gly Lys Lys Glu Thr Leu Val Glu Asp Lys Leu
1640 1645 1650

Ser Leu Leu Asn Ser Asn Trp Ile Ala Val Thr Ser Arg Ala Glu
1655 1660 1665

Glu Trp Leu Asn Leu Leu Leu Glu Tyr Gln Lys His Met Glu Thr
1670 1675 1680

Phe Asp Gln Asn Val Asp His Ile Thr Lys Trp Ile Ile Gln Ala
1685 1690 1695

Asp Thr Leu Leu Asp Glu Ser Glu Lys Lys Lys Pro Gln Gln Lys
1700 1705 1710

Glu Asp Val Leu Lys Arg Leu Lys Ala Glu Leu Asn Asp Ile Arg
1715 1720 1725

Pro Lys Val Asp Ser Thr Arg Asp Gln Ala Ala Asn Leu Met Ala
1730 1735 1740

Asn Arg Gly Asp His Cys Arg Lys Leu Val Glu Pro Gln Ile Ser
1745 1750 1755

Glu Leu Asn His Arg Phe Ala Ala Ile Ser His Arg Ile Lys Thr
1760 1765 1770

Gly Lys Ala Ser Ile Pro Leu Lys Glu Leu Glu Gln Phe Asn Ser
1775 1780 1785

Asp Ile Gln Lys Leu Leu Glu Pro Leu Glu Ala Glu Ile Gln Gln
1790 1795 1800

Gly Val Asn Leu Lys Glu Glu Asp Phe Asn Lys Asp Met Asn Glu
1805 1810 1815

Asp Asn Glu Gly Thr Val Lys Glu Leu Leu Gln Arg Gly Asp Asn
1820 1825 1830

Leu Gln Gln Arg Ile Thr Asp Glu Arg Lys Arg Glu Glu Ile Lys
1835 1840 1845

Ile Lys Gln Gln Leu Leu Gln Thr Lys His Asn Ala Leu Lys Asp
1850 1855 1860

Leu Arg Ser Gln Arg Arg Lys Lys Ala Leu Glu Ile Ser His Gln
1865 1870 1875

Trp Tyr Gln Tyr Lys Arg Gln Ala Asp Asp Leu Leu Lys Cys Leu
1880 1885 1890

Asp Asp Ile Glu Lys Lys Leu Ala Ser Leu Pro Glu Pro Arg Asp
1895 1900 1905

Glu Arg Lys Ile Lys Glu Ile Asp Arg Glu Leu Gln Lys Lys Lys
1910 1915 1920

Glu Glu Leu Asn Ala Val Arg Arg Gln Ala Glu Gly Leu Ser Glu
1925 1930 1935

Asp Gly Ala Ala Met Ala Val Glu Pro Thr Gln Ile Gln Leu Ser
1940 1945 1950

Lys Arg Trp Arg Glu Ile Glu Ser Lys Phe Ala Gln Phe Arg Arg
1955 1960 1965

Leu Asn Phe Ala Gln Ile His Thr Val Arg Glu Glu Thr Met Met
1970 1975 1980

Val Met Thr Glu Asp Met Pro Leu Glu Ile Ser Tyr Val Pro Ser

1985 1990 1995

Thr Tyr Leu Thr Glu Ile Thr His Val Ser Gln Ala Leu Leu Glu
2000 2005 2010

Val Glu Gln Leu Leu Asn Ala Pro Asp Leu Cys Ala Lys Asp Phe
2015 2020 2025

Glu Asp Leu Phe Lys Gln Glu Glu Ser Leu Lys Asn Ile Lys Asp
2030 2035 2040

Ser Leu Gln Gln Ser Ser Gly Arg Ile Asp Ile Ile His Ser Lys
2045 2050 2055

Lys Thr Ala Ala Leu Gln Ser Ala Thr Pro Val Glu Arg Val Lys
2060 2065 2070

Leu Gln Glu Ala Leu Ser Gln Leu Asp Phe Gln Trp Glu Lys Val
2075 2080 2085

Asn Lys Met Tyr Lys Asp Arg Gln Gly Arg Phe Asp Arg Ser Val
2090 2095 2100

Glu Lys Trp Arg Arg Phe His Tyr Asp Ile Lys Ile Phe Asn Gln
2105 2110 2115

Trp Leu Thr Glu Ala Glu Gln Phe Leu Arg Lys Thr Gln Ile Pro
2120 2125 2130

Glu Asn Trp Glu His Ala Lys Tyr Lys Trp Tyr Leu Lys Glu Leu
2135 2140 2145

Gln Asp Gly Ile Gly Gln Arg Gln Thr Val Val Arg Thr Leu Asn
2150 2155 2160

Ala Thr Gly Glu Glu Ile Ile Gln Gln Ser Ser Lys Thr Asp Ala
2165 2170 2175

Ser Ile Leu Gln Glu Lys Leu Gly Ser Leu Asn Leu Arg Trp Gln
2180 2185 2190

Glu Val Cys Lys Gln Leu Ser Asp Arg Lys Lys Arg Leu Glu Glu
2195 2200 2205

Gln Lys Asn Ile Leu Ser Glu Phe Gln Arg Asp Leu Asn Glu Phe
2210 2215 2220

Val Leu Trp Leu Glu Glu Ala Asp Asn Ile Ala Ser Ile Pro Leu
2225 2230 2235

Glu Pro Gly Lys Glu Gln Gln Leu Lys Glu Lys Leu Glu Gln Val
2240 2245 2250

Lys Leu Leu Val Glu Glu Leu Pro Leu Arg Gln Gly Ile Leu Lys
2255 2260 2265

Gln Leu Asn Glu Thr Gly Gly Pro Val Leu Val Ser Ala Pro Ile
2270 2275 2280

Ser Pro Glu Glu Gln Asp Lys Leu Glu Asn Lys Leu Lys Gln Thr
2285 2290 2295

Asn Leu Gln Trp Ile Lys Val Ser Arg Ala Leu Pro Glu Lys Gln
2300 2305 2310

Gly Glu Ile Glu Ala Gln Ile Lys Asp Leu Gly Gln Leu Glu Lys
2315 2320 2325

Lys Leu Glu Asp Leu Glu Glu Gln Leu Asn His Leu Leu Leu Trp
2330 2335 2340

Leu Ser Pro Ile Arg Asn Gln Leu Glu Ile Tyr Asn Gln Pro Asn
2345 2350 2355

Gln Glu Gly Pro Phe Asp Val Gln Glu Thr Glu Ile Ala Val Gln
2360 2365 2370

Ala Lys Gln Pro Asp Val Glu Glu Ile Leu Ser Lys Gly Gln His
2375 2380 2385

Leu Tyr Lys Glu Lys Pro Ala Thr Gln Pro Val Lys Arg Lys Leu
2390 2395 2400

Glu Asp Leu Ser Ser Glu Trp Lys Ala Val Asn Arg Leu Leu Gln
2405 2410 2415

Glu Leu Arg Ala Lys Gln Pro Asp Leu Ala Pro Gly Leu Thr Thr
2420 2425 2430

Ile Gly Ala Ser Pro Thr Gln Thr Val Thr Leu Val Thr Gln Pro
2435 2440 2445

Val Val Thr Lys Glu Thr Ala Ile Ser Lys Leu Glu Met Pro Ser
2450 2455 2460

Ser Leu Met Leu Glu Val Pro Ala Leu Ala Asp Phe Asn Arg Ala
2465 2470 2475

Trp Thr Glu Leu Thr Asp Trp Leu Ser Leu Leu Asp Gln Val Ile
2480 2485 2490

Lys Ser Gln Arg Val Met Val Gly Asp Leu Glu Asp Ile Asn Glu
2495 2500 2505

Met Ile Ile Lys Gln Lys Ala Thr Met Gln Asp Leu Glu Gln Arg
2510 2515 2520

Arg Pro Gln Leu Glu Glu Leu Ile Thr Ala Ala Gln Asn Leu Lys
2525 2530 2535

Asn Lys Thr Ser Asn Gln Glu Ala Arg Thr Ile Ile Thr Asp Arg
2540 2545 2550

Ile Glu Arg Ile Gln Asn Gln Trp Asp Glu Val Gln Glu His Leu
2555 2560 2565

Gln Asn Arg Arg Gln Gln Leu Asn Glu Met Leu Lys Asp Ser Thr
2570 2575 2580

Gln Trp Leu Glu Ala Lys Glu Glu Ala Glu Gln Val Leu Gly Gln
2585 2590 2595

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

2600 2605 2610

Asp Ala Ile Gln Lys Lys Ile Thr Glu Thr Lys Gln Leu Ala Lys
2615 2620 2625

Asp Leu Arg Gln Trp Gln Thr Asn Val Asp Val Ala Asn Asp Leu
2630 2635 2640

Ala Leu Lys Leu Leu Arg Asp Tyr Ser Ala Asp Asp Thr Arg Lys
2645 2650 2655

Val His Met Ile Thr Glu Asn Ile Asn Ala Ser Trp Arg Ser Ile
2660 2665 2670

His Lys Arg Val Ser Glu Arg Glu Ala Ala Leu Glu Glu Thr His
2675 2680 2685

Arg Leu Leu Gln Gln Phe Pro Leu Asp Leu Glu Lys Phe Leu Ala
2690 2695 2700

Trp Leu Thr Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala
2705 2710 2715

Thr Arg Lys Glu Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu
2720 2725 2730

Leu Met Lys Gln Trp Gln Asp Leu Gln Gly Glu Ile Glu Ala His
2735 2740 2745

Thr Asp Val Tyr His Asn Leu Asp Glu Asn Ser Gln Lys Ile Leu
2750 2755 2760

Arg Ser Leu Glu Gly Ser Asp Asp Ala Val Leu Leu Gln Arg Arg
2765 2770 2775

Leu Asp Asn Met Asn Phe Lys Trp Ser Glu Leu Arg Lys Lys Ser
2780 2785 2790

Leu Asn Ile Arg Ser His Leu Glu Ala Ser Ser Asp Gln Trp Lys
2795 2800 2805

Arg Leu His Leu Ser Leu Gln Glu Leu Leu Val Trp Leu Gln Leu
2810 2815 2820

Lys Asp Asp Glu Leu Ser Arg Gln Ala Pro Ile Gly Gly Asp Phe
2825 2830 2835

Pro Ala Val Gln Lys Gln Asn Asp Val His Arg Ala Phe Lys Arg
2840 2845 2850

Glu Leu Lys Thr Lys Glu Pro Val Ile Met Ser Thr Leu Glu Thr
2855 2860 2865

Val Arg Ile Phe Leu Thr Glu Gln Pro Leu Glu Gly Leu Glu Lys
2870 2875 2880

Leu Tyr Gln Glu Pro Arg Glu Leu Pro Pro Glu Glu Arg Ala Gln
2885 2890 2895

Asn Val Thr Arg Leu Leu Arg Lys Gln Ala Glu Glu Val Asn Thr
2900 2905 2910

Glu Trp Glu Lys Leu Asn Leu His Ser Ala Asp Trp Gln Arg Lys
2915 2920 2925

Ile Asp Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln Glu Ala Thr
2930 2935 2940

Asp Glu Leu Asp Leu Lys Leu Arg Gln Ala Glu Val Ile Lys Gly
2945 2950 2955

Ser Trp Gln Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln Asp
2960 2965 2970

His Leu Glu Lys Val Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu
2975 2980 2985

Lys Glu Asn Val Ser His Val Asn Asp Leu Ala Arg Gln Leu Thr
2990 2995 3000

Thr Leu Gly Ile Gln Leu Ser Pro Tyr Asn Leu Ser Thr Leu Glu
3005 3010 3015

Asp Leu Asn Thr Arg Trp Lys Leu Leu Gln Val Ala Val Glu Asp
3020 3025 3030

Arg Val Arg Gln Leu His Glu Ala His Arg Asp Phe Gly Pro Ala
3035 3040 3045

Ser Gln His Phe Leu Ser Thr Ser Val Gln Gly Pro Trp Glu Arg
3050 3055 3060

Ala Ile Ser Pro Asn Lys Val Pro Tyr Tyr Ile Asn His Glu Thr
3065 3070 3075

Gln Thr Thr Cys Trp Asp His Pro Lys Met Thr Glu Leu Tyr Gln
3080 3085 3090

Ser Leu Ala Asp Leu Asn Asn Val Arg Phe Ser Ala Tyr Arg Thr
3095 3100 3105

Ala Met Lys Leu Arg Arg Leu Gln Lys Ala Leu Cys Leu Asp Leu
3110 3115 3120

Leu Ser Leu Ser Ala Ala Cys Asp Ala Leu Asp Gln His Asn Leu
3125 3130 3135

Lys Gln Asn Asp Gln Pro Met Asp Ile Leu Gln Ile Ile Asn Cys
3140 3145 3150

Leu Thr Thr Ile Tyr Asp Arg Leu Glu Gln Glu His Asn Asn Leu
3155 3160 3165

Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn Trp Leu Leu
3170 3175 3180

Asn Val Tyr Asp Thr Gly Arg Thr Gly Arg Ile Arg Val Leu Ser
3185 3190 3195

Phe Lys Thr Gly Ile Ile Ser Leu Cys Lys Ala His Leu Glu Asp
3200 3205 3210

Lys Tyr Arg Tyr Leu Phe Lys Gln Val Ala Ser Ser Thr Gly Phe

3215 3220 3225

Cys Asp Gln Arg Arg Leu Gly Leu Leu Leu His Asp Ser Ile Gln
3230 3235 3240

Ile Pro Arg Gln Leu Gly Glu Val Ala Ser Phe Gly Gly Ser Asn
3245 3250 3255

Ile Glu Pro Ser Val Arg Ser Cys Phe Gln Phe Ala Asn Asn Lys
3260 3265 3270

Pro Glu Ile Glu Ala Ala Leu Phe Leu Asp Trp Met Arg Leu Glu
3275 3280 3285

Pro Gln Ser Met Val Trp Leu Pro Val Leu His Arg Val Ala Ala
3290 3295 3300

Ala Glu Thr Ala Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu
3305 3310 3315

Cys Pro Ile Ile Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn
3320 3325 3330

Tyr Asp Ile Cys Gln Ser Cys Phe Phe Ser Gly Arg Val Ala Lys
3335 3340 3345

Gly His Lys Met His Tyr Pro Met Val Glu Tyr Cys Thr Pro Thr
3350 3355 3360

Thr Ser Gly Glu Asp Val Arg Asp Phe Ala Lys Val Leu Lys Asn
3365 3370 3375

Lys Phe Arg Thr Lys Arg Tyr Phe Ala Lys His Pro Arg Met Gly
3380 3385 3390

Tyr Leu Pro Val Gln Thr Val Leu Glu Gly Asp Asn Met Glu Thr
3395 3400 3405

Pro Val Thr Leu Ile Asn Phe Trp Pro Val Asp Ser Ala Pro Ala
3410 3415 3420

Ser Ser Pro Gln Leu Ser His Asp Asp Thr His Ser Arg Ile Glu
3425 3430 3435

His Tyr Ala Ser Arg Leu Ala Glu Met Glu Asn Ser Asn Gly Ser
3440 3445 3450

Tyr Leu Asn Asp Ser Ile Ser Pro Asn Glu Ser Ile Asp Asp Glu
3455 3460 3465

His Leu Leu Ile Gln His Tyr Cys Gln Ser Leu Asn Gln Asp Ser
3470 3475 3480

Pro Leu Ser Gln Pro Arg Ser Pro Ala Gln Ile Leu Ile Ser Leu
3485 3490 3495

Glu Ser Glu Glu Arg Gly Glu Leu Glu Arg Ile Leu Ala Asp Leu
3500 3505 3510

Glu Glu Glu Asn Arg Asn Leu Gln Ala Glu Tyr Asp Arg Leu Lys
3515 3520 3525

Gln Gln His Glu His Lys Gly Leu Ser Pro Leu Pro Ser Pro Pro
3530 3535 3540

Glu Met Met Pro Thr Ser Pro Gln Ser Pro Arg Asp Ala Glu Leu
3545 3550 3555

Ile Ala Glu Ala Lys Leu Leu Arg Gln His Lys Gly Arg Leu Glu
3560 3565 3570

Ala Arg Met Gln Ile Leu Glu Asp His Asn Lys Gln Leu Glu Ser
3575 3580 3585

Gln Leu His Arg Leu Arg Gln Leu Leu Glu Gln Pro Gln Ala Glu
3590 3595 3600

Ala Lys Val Asn Gly Thr Thr Val Ser Ser Pro Ser Thr Ser Leu
3605 3610 3615

Gln Arg Ser Asp Ser Ser Gln Pro Met Leu Leu Arg Val Val Gly
3620 3625 3630

Ser Gln Thr Ser Asp Ser Met Gly Glu Glu Asp Leu Leu Ser Pro
3635 3640 3645

Pro Gln Asp Thr Ser Thr Gly Leu Glu Glu Val Met Glu Gln Leu
3650 3655 3660

Asn Asn Ser Phe Pro Ser Ser Arg Gly Arg Asn Thr Pro Gly Lys
3665 3670 3675

Pro Met Arg Glu Asp Thr Met
3680 3685

<210> 65
<211> 19841
<212> DNA
<213> Loài người
<400> 65
tttcctactt gaagcctgac gtagtaaaga tcgggggagg gttagacaga tacagtggtc
60
cccaaccatt ttggcaccag ggactggtct tatggaagac agtttttcca cagactgttg
120
gggggatggt tttgggatga aaccgttctg cctctgatca tcaggtgtta gattctaata
180
aggagcgcac acctagatcc ctcgcatgca tagttcatgg tggggttcgc actcttacga
240
ggattgaatg gtgcgctgct ccggtaggag gctgggctca ggctgtaatg cctgctcgcc
300
caccactcac ctcctgctgc atggcctggt tcctaacagg ccacggacca ctactgttcc
360
atggcccgga ggttgaggac cccgagatac aggacaattc tgtggcaagc aggactgtcc
420
cctcgccaaa gatgggacat cgaggctcct tggagcaccc tgtggccacc ttgcagcagc
480
ctctgtttcc ccatgtttcc atgacctggt gtccatctgt cttccccagt ttgggagctt
540
ctctccgagg aggacctggg cctggtgtgg cacctgctgt gtgagcgggg ccatgtccaa
600
cggccttctt ggggactttg ggtcggggag aagttctgcc tgggttttac tgccttctcc
660
caaccccaca ctgtctcccc tggcagcggt tgatcgacaa gaccaaggtg acatatctga
720

agtggctgcc tgagtcggag agcctgttcc tggcatcaca cgccagtggc cacctgtacc
780
tgtacaacgt cagccacccc tgcgcctcgg ccccgcccca gtacagcctg ctgaagcagg
840
gcgagggctt ctctgtctat gctgccaaga gcaaggcacc ccgcaacccg ctggccaagt
900
gggcggtggg tgaggggccc ctcaacgagt tcgccttctc gcccgatggc cggcacctgg
960
cctgtgtgag ccaggatggc tgcctgcgcg tcttccactt cgactccatg ctcctgcgtg
1020
ggctcatgaa gagctacttt gggggcctgc tgtgtgtgtg ctggagccct gacggccgct
1080
acgtggtgac gggtggcgaa gatgacctgg tcaccgtgtg gtccttcacc gagggccgcg
1140
tggtggctcg aggccatggc cacaagtcct gggtcaacgc tgtggccttt gacccctaca
1200
ccacaagggc agaggaggcg gcgacagcag ccggtgctga tggggagcgg agcggcgaag
1260
aggaggagga ggagcccgag gctgcgggca caggctcggc cgggggcgcc ccgctctctc
1320
cactgcccaa ggctggctcc attacttacc gctttggctc ggcgggccag gacacgcagt
1380
tctgcctgtg ggacctcact gaagacgtgc tctacccgca ccccccccctg gcccgcaccc
1440
gcaccctccc tggcacacct ggcaccacgc caccggccgc cagcagctcg aggggtggcg
1500
agcctggccc aggcccccctg cctcgctcgc tgtcccgctc caacagtctc ccgcacccag
1560
ctggcggggg caaggcgggc ggcccgggtg tggcggcaga gcctggcaca ccattcagca
1620
ttggccgctt cgccacgctc acactgcagg agcggcggga ccgggggggca gagaaggagc
1680
acaagcgcta ccacagcctg ggcaacatca gccggggtgg cagtggcggc agtggcagtg
1740
gtggggagaa gcccagcggc cctgttcccc gcagccgcct ggaccccgcc aaggtgctgg
1800
gcactgcgct gtgcccgcgc atccacgagg tgcccctgct ggagcccctt gtgtgcaaga
1860
agatcgccca ggagcggctc acagtcctcc tgttcctgga ggactgcatc atcactgcct
1920
gccaggaggg cctcatctgc acctgggccc ggccgggcaa ggcggtgagt ggccccacac
1980
cagcctgccg gggacctggc aggacctttc gtgggaagag gcaggcattg gcagagagag
2040

ggctttgttg ctgtcacagc ctctggctcc gtggggtgag gggaagccag ggaaatctta
2100
gtgtctcagt acaagacctc tcagatcctt agagtgaggg ggtctagccc taggcagcag
2160
gcagcagaaa gaggggtggg tgtgagagcc agctaggaat tggggcatcc aaggctggcc
2220
gtctgaaggg cagcagatgg gccccacatg gccaggtctt actgcctgtc actcgaacca
2280
gaatctattt ctgttgaaca tctgtttttt aaatcgtgaa actttttga gtacttcagg
2340
ccaaaactag gggcgagctc aagcctgtgg gcatggctgc cagcctgggt ctgggactca
2400
ggatctgagc ctcctgctga aggcacaggc tgggaatccc aggcctgggt tccagtccca
2460
ctccctctgt gaccctggac aagtcactgc cccctctgac ctccaactca tcacctctta
2520
gaacagagcc tgtaggatgg gcagtgggtg gatgtgcttg cctcctgggt gggctgtggc
2580
gttgggaagg tcatagtagg cgaatcaggc ctggcatctt gtaagttcgg agctcgtctt
2640
gggtgtctca gcttcttagg gcttggactc agttgcccag ggtcctggag gccgtggctt
2700
ggttcctcag atcctcagtt ttggaatcgt agagtcctga gtccctagaa cttgagagca
2760
cagtctgagt gactcagagg caagagtggt gggatttggg gagtctggtt gagtcctaaa
2820
agagacccct ctgtctccgt agttcacaga cgaggagacc gaggcccaga caggggaagg
2880
aagttggccc aggtcaccca gcaagtcagt ggtagaggta ggactgtccc tgagttcttt
2940
ccccagcacc tcagggtccc tcccaagtta gaagggagct ccagtttccc cctcccctcc
3000
caccettacc cttaccccat ggtctcactc aggatccgcc aaggactttg attattgcgt
3060
gaaagtgctg actgccagga caggaagcta gctaagatgc aagttcccag cctagagcag
3120
tggcctctgg ggggtctagg gcggacccaa gggcaaggcc agggtggcag cagctttggg
3180
gactctgggc tggctccctc cccttgacac tggctgaagc ccaggtggtc tctaacccct
3240
cccatctctc cctctcatct tccccagggc atctcctccc aaccaggcaa ctccccgagt
3300
ggcacagtgg tgtgaagcca tggatatcgg gccccccaa ccccatgccc ccagcctcct
3360

agccataacc ctccctgctg acctcacaga tcaacgtatt aacaagacta accatgatgg
3420
atggactgct ccagtccccc cacctgcaca aaatttgggg gccccccaga ctggcccgga
3480
cacgggcgat gtaatagccc ttgtggcctc agccttgtcc cccacccact gccaagtaca
3540
atgacctctt cctctgaaac atcagtgtta ccctcatccc tgtccccagc atgtgactgg
3600
tcactcctgg ggagagactc cccgcccctg ccacaagagc cccaggtctg cagtgtgccc
3660
ctcagttgag tgggcagggc cgggggtggt ccagccctcg cccggccccc accccagctg
3720
cccttgctat tgtctgtgct tttgaagagt gttaaattat ggaagcccct caggttcctc
3780
cctgtcccgc aggacctctt atttatacta aagttccctg ttttctcagc gggtctgtcc
3840
ccttcggagg agatgatgta gaggacctgt gtgtgtactc tgtggttcta ggcagtccgc
3900
tttccccaga ggaggagtgc aggcctgctc ccagcccagc gcctcccacc ccttttcata
3960
gcaggaaaag ccggagccca gggagggaac ggacctgcga gtcacacaac tggtgaccca
4020
caccagcggc tggagcagga ccctcttggg gagaagagca tcctgcccgc agccagggcc
4080
cctcatcaaa gtcctcggtg ttttttaaat tatcagaact gcccaggacc acgtttccca
4140
ggccctgccc agctgggact cctcggtcct tgcctcctag tttctcaggc ctggccctct
4200
caaggcccag gcaccccagg ccggttggag gccccgactt ccactctgga gaaccgtcca
4260
ccctggaaag aagagctcag attcctcttg gctctcggag ccgcagggag tgtgtcttcc
4320
cgcgccaccc tccacccccc gaaatgtttc tgtttctaat cccagcctgg gcaggaatgt
4380
ggctccccgg ccaggggcca aggagctatt ttggggtctc gtttgcccag ggagggcttg
4440
gctccaccac tttcctcccc cagcctttgg gcagcaggtc acccctgttc aggctctgag
4500
ggtgccccct cctggtcctg tcctcaccac cccttcccca cctcctggga aaaaaaaaaa
4560
aaaaaaaaaa aaaagctggt ataaagcaga gagcctgagg gctaaattta actgtccgag
4620
tcggaatcca tctctgagtc acccaagaag ctgccctggc ctcccgtccc cttcccaggc
4680

ctcaacccct ttctcccacc cagccccaac ccccagccct cacccctag cccccagttc
4740
tggagcttgt cgggagcaag ggggtggttg ctactgggtc actcagcctc aattggccct
4800
gtttcagcaa tgggcaggtt cttcttgaaa ttcatcacac ctgtggcttc ctctgtgctc
4860
tacctttta ttggggtgac agtgtgacag ctgagattct ccatgcattc cccctactct
4920
agcactgaag ggttctgaag ggccctggaa ggagggagct tgggggctg gcttgtgagg
4980
ggttaaggct gggaggcggg agggggctg gaccaagggg tggggagaag gggaggaggc
5040
ctcggccggc cgcagagaga agtggccaga gaggcccagg ggacagccag ggacaggcag
5100
acatgcagcc agggctccag ggcctggaca ggggctgcca ggccctgtga caggaggacc
5160
ccgagccccc ggcccgggga ggggccatgg tgctgcctgt ccaacatgtc agccgaggtg
5220
cggctgaggc ggctccagca gctggtgttg gacccgggct tcctggggct ggagcccctg
5280
ctcgaccttc tcctgggcgt ccaccaggag ctgggcgcct ccgaactggc ccaggacaag
5340
tacgtggccg acttcttgca gtggggtgag tgcctaccct cggggctcct gcagatgggg
5400
tgggggtggg gcaggagaca ggtctgggca cagaggcctg gctgttgggg gggcaggatg
5460
gcaggatggg catggggaga tcctcccatc ctggggctca gagtgtggac ctgggccctg
5520
gggcaacatt tctctgtcct atgccaccac tctggagggg cagagtaagg tcagcagagg
5580
ctagggtggc tgtgactcag agccatggct taggagtcac agcaggctag gctgccaaca
5640
gcctcccatg gcctctctgc accccgcctc agggtcaggg tcagggtcat gctgggagct
5700
ccctctccta ggaccctccc cccaaaagtg ggctctatgg ccctctcccc tggtttcctg
5760
tggcctgggg caagccagga gggccagcat ggggcagctg ccaggggcgc agccgacagg
5820
caggtgttcg gcgccagcct ctccagctgc cccaacaggt gcccaggcac tgggagggcg
5880
gtgactcacg cgggccctgt gggagaacca gctttgcaga caggcgccac cagtgccccc
5940
tcctctgcga tccaggaggg acaactttgg gttcttctgg gtgtgtctcc ttcttttgta
6000

ggttctgcac ccacccccac ccccagcccc aaagtctcgg ttcctatgag ccgtgtgggt
6060
cagccaccat tcccgccacc ccgggtccct gcgtccttta gttctcctgg cccagggcct
6120
ccaaccttcc agctgtccca caaaacccct tcttgcaagg gctttccagg gcctggggcc
6180
agggctggaa ggaggatgct tccgcttctg ccagctgcct tgtctgccca cctcctcccc
6240
aagcccagga ctcgggctca ctggtcactg gtttctttca ttcccagcac cctgcccctc
6300
tggccctcat atgtctggcc ctcagtgact ggtgtttggt ttttggcctg tgtgtaacaa
6360
actgtgtgtg acacttgttt cctgtttctc cgccttcccc tgcttcctct tgtgtccatc
6420
tctttctgac ccaggcctgg ttcctttccc tcctcctccc atttcacaga tgggaaggtg
6480
gaggccaaga agggccaggc cattcagcct ctggaaaaac cttctcccaa cctcccacag
6540
cccctaatga ctctcctggc ctccctttag tagaggatga agttgggttg gcagggtaaa
6600
ctgagaccgg gtggggtagg ggtctggcgc tcccgggagg agcactcctt ttgtggcccg
6660
agctgcatct cgcggccccc cccctgccag gcctggggcg ggggaggggg ccagggttcc
6720
tgctgcctta aaagggctca atgtcttggc tctctcctcc ctcccccgtc ctcagccctg
6780
gctggttcgt ccctgctggc ccactctccc ggaacccccc ggaacccctc tctttcctcc
6840
agaacccact gtctcctctc cttccctccc ctcccatacc catccctctc tccatcctgc
6900
ctccacttct tccacccccg ggagtccagg cctccctgtc cccacagtcc ctgagccaca
6960
agcctccacc ccagctggtc ccccacccag gctgcccagt ttaacattcc tagtcatagg
7020
accttgactt ctgagaggcc tgattgtcat ctgtaaataa ggggtaggac taaagcactc
7080
ctcctggagg actgagagat gggctggacc ggagcacttg agtctgggat atgtgaccat
7140
gctacctttg tctccctgtc ctgttccttc ccccagcccc aaatccaggg ttttccaaag
7200
tgtggttcaa gaaccacctg catctgaatc tagaggtact ggatacaacc ccacgtctgg
7260
gccgttaccc aggacattct acatgagaac gtgggggtgg ggccctggct gcacctgaac
7320

tgtcacctgg agtcagggtg gaaggtggaa gaactgggtc ttatttcctt ctccccttgt
7380
tctttagggt ctgtccttct gcagactccg ttaccccacc ctaaccatcc tgcacaccct
7440
tggagccctc tgggccaatg ccctgtcccg caaagggctt ctcaggcatc tcacctctat
7500
gggagggcat ttttggcccc cagaacctta cacggtgttt atgtggggaa gcccctggga
7560
agcagacagt cctagggtga agctgagagg cagagagaag gggagacaga cagagggtgg
7620
ggctttcccc cttgtctcca gtgccctttc tggtgaccct cggttctttt cccccaccac
7680
ccccccagcg gagcccatcg tggtgaggct taaggaggtc cgactgcaga gggacgactt
7740
cgagattctg aaggtgatcg gacgcggggc gttcagcgag gtaagccgaa ccgggcggga
7800
gcctgacttg actcgtggtg ggcggggcat aggggttggg gcggggcctt agaaattgat
7860
gaatgaccga gccttagaac ctagggctgg gctggaggcg gggcttggga ccaatgggcg
7920
tggtgtggca ggtggggcgg ggccacggct gggtgcagaa gcgggtggag ttgggtctgg
7980
gcgagccctt ttgttttccc gccgtctcca ctctgtctca ctatctcgac ctcaggtagc
8040
ggtagtgaag atgaagcaga cgggccaggt gtatgccatg aagatcatga acaagtggga
8100
catgctgaag aggggcgagg tgaggggctg ggcggacgtg gggggctttg aggatccgcg
8160
ccccgtctcc ggctgcagct cctccgggtg ccctgcaggt gtcgtgcttc cgtgaggaga
8220
gggacgtgtt ggtgaatggg gaccggcggt ggatcacgca gctgcacttc gccttccagg
8280
atgagaacta cctggtgagc tccgggccgg ggtgactagg aagagggaca agagcccgtg
8340
ctgtcactgg acgaggaggt ggggagagga agctctagga ttgggggtgc tgcccggaaa
8400
cgtctgtggg aaagtctgtg tgcggtaaga gggtgtgtca ggtggatgag gggccttccc
8460
tatctgagac ggggatggtg tccttcactg cccgtttctg gggtgatctg ggggactctt
8520
ataaagatgt ctctgttgcg gggggtctct tacctggaat gggataggtc ttcaggaatt
8580
ctaacggggc cactgcctag ggaaggagtg tctgggacct attctctggg tgttgggtgg
8640

cctctgggtt ctctttccca gaacatctca gggggagtga atctgcccag tgacatccca
8700
ggaaagtttt tttgtttgtg ttttttttg aggggcgggg gcgggggccg caggtggtct
8760
ctgatttggc ccggcagatc tctatggtta tctctgggct ggggctgcag gtctctgccc
8820
aaggatgggg tgtctctggg aggggttgtc ccagccatcc gtgatggatc agggcctcag
8880
gggactacca accacccatg acgaacccct tctcagtacc tggtcatgga gtattacgtg
8940
ggcggggacc tgctgacact gctgagcaag tttggggagc ggattccggc cgagatggcg
9000
cgcttctacc tggcggagat tgtcatggcc atagactcgg tgcaccggct tggctacgtg
9060
cacaggtggg tgcagcatgg ccgaggggat agcaagcttg ttccctggcc gggttcttgg
9120
aaggtcagag cccagagagg ccagggcctg gagagggacc ttcttggttg gggcccaccg
9180
gggggtgcct gggagtaggg gtcagaactg tagaagccct acaggggcgg aacccgagga
9240
agtggggtcc caggtggcac tgcccggagg ggcggagcct ggtgggacca cagaagggag
9300
gttcatttat cccacccttc tcttttcctc cgtgcaggga catcaaaccc gacaacatcc
9360
tgctggaccg ctgtggccac atccgcctgg ccgacttcgg ctcttgcctc aagctgcggg
9420
cagatggaac ggtgagccag tgccctggcc acagagcaac tggggctgct gatgagggat
9480
ggaaggcaca gagtgtggga gcgggactgg atttggaggg gaaaagaggt ggtgtgaccc
9540
aggcttaagt gtgcatctgt gtggcggagt attagaccag gcagagggag gggctaagca
9600
tttggggagt ggttggaagg agggcccaga gctggtgggc ccagaggggt gggcccaagc
9660
ctcgctctgc tccttttggt ccaggtgcgg tcgctggtgg ctgtgggcac cccagactac
9720
ctgtcccccg agatcctgca ggctgtgggc ggtgggcctg ggacaggcag ctacgggccc
9780
gagtgtgact ggtgggcgct gggtgtattc gcctatgaaa tgttctatgg gcagacgccc
9840
ttctacgcgg attccacggc ggagacctat ggcaagatcg tccactacaa ggtgagcacg
9900
gccgcaggga gacctggcct ctcccggtag gcgctcccag gctatcgcct cctctccctc
9960

tgagcaggag cacctctctc tgccgctggt ggacgaaggg gtccctgagg aggctcgaga
10020
cttcattcag cggttgctgt gtcccccgga gacacggctg ggccggggtg gagcaggcga
10080
cttccggaca catcccttct tctttggcct cgactgggat ggtctccggg acagcgtgcc
10140
ccccttttaca ccggatttcg aaggtgccac cgacacatgc aacttcgact tggtggagga
10200
cgggctcact gccatggtga gcgggggcgg ggtaggtacc tgtggcccct gctcggctgc
10260
gggaacctcc ccatgctccc tccataaagt tggagtaagg acagtgccta ccttctgggg
10320
tcctgaatca ctcattcccc agagcacctg ctctgtgccc atctactact gaggacccag
10380
cagtgaccta gacttacagt ccagtggggg aacacagagc agtcttcaga cagtaaggcc
10440
ccagagtgat cagggctgag acaatggagt gcaggggtg ggggactcct gactcagcaa
10500
ggaaggtcct ggagggcttt ctggagtggg gagctatctg agctgagact tggagggatg
10560
agaagcagga gaggactcct cctcccttag gccgtctctc ttcaccgtgt aacaagctgt
10620
catggcatgc ttgctcggct ctgggtgccc ttttgctgaa caatactggg gatccagcac
10680
ggaccagatg agctctggtc cctgccctca tccagttgca gtctagagaa ttagagaatt
10740
atggagagtg tggcaggtgc cctgaaggga agcaacagga tacaagaaaa aatgatgggg
10800
ccaggcacgg tggctcacgc ctgtaacccc agcaatttgg caggccgaag tgggtggatt
10860
gcttgagccc aggagttcga gaccagcctg ggcaatgtgg tgagaccccc gtctctacaa
10920
aaatgtttta aaaattggtt gggcgtggtg gcgcatgcct gtatactcag ctactagggt
10980
ggccgacgtg ggcttgagcc caggaggtca aggctgcagt gagctgtgat tgtgccactg
11040
cactccagcc tgggcaacgg agagagactc tgtctcaaaa ataagataaa ctgaaattaa
11100
aaaataggct gggctggccg ggcgtggtgg ctcacgcctg taatctcagc actttgggag
11160
gccgaggcgg gtggatcacg aggtcaggag atcgagacca tcttggctaa cacggtgaaa
11220
ccccatctct cctaaaaata caaaaaatta gccaggcgtg gtggcgggcg cctgtagtcc
11280

cagctactca ggaggctgag gcaggagaat ggcgtgaacc cgggaggcag agtttgcagt
11340
gagccgagat cgtgccactg cactccagcc tgggcgacag agcgagactc tgtctcagaa
11400
aaaaaaaaaa aaaaaaaaaa aaataggctg gaccgcggcc gggcgctgtg gctcatgcct
11460
gtaatcccag cactttggga gtccaaggcc ggtgggtcat gagatcagga gttttgagac
11520
taggctggcc aacacggtga aaccccgtct ctactaaaaa tacaagaaaa ttagctgggt
11580
gtggtctcgg gtgcctgtaa ttccagttac tggggaagct gaggcaggag aattgcttga
11640
acctgggagg cagagtttgc agtgagccaa gatcatgcca ctacactcca gtctgggtga
11700
cagagtgaga ctctgtctca aaaaaaaaaa aaaaaaaaag ggttgggcaa ggtggttcac
11760
gcctgtaatc ccagaacttt gggaggctga ggcaggcaga tcactggaag tcaggagttc
11820
aagaccagcc tggccaacat ggtgaaaccc tgtgtctact aaaaatacaa aatttagcca
11880
ggcttggtgg cgtatgcctg taatgccagc tactcaggag gctgaggcag gagaatcgct
11940
tgattgaacc tgggaggcag agtttgcagt gggctggggt tgtgccactg cactctaggc
12000
tgggagacag caagactcca tctaaaaaaa aaaaacagaa ctgggctggg cacagtggct
12060
tatatttgta atcccagcac tttgggaggc tgaggttgga ggactgcttg agcccagagt
12120
ttgggactac aacagctgag gtaggcggat cacttgaggt cagaagatgg agaccagcct
12180
ggccagcgtg gcgaaacccc gtctctacca aaaatataaa aaattagcca ggcgtggtag
12240
agggcgcctg taatctcagc tactcaggac gctgaggcag gagaatcgcc tgaacctggg
12300
aggcggaggt tgcagtgagc tgagattgca ccactgcact ccagcctggg taacagagcg
12360
agactccgta tcaaagaaaa agaaaaaaga aaaaatgctg gaggggccac tttagataag
12420
ccctgagttg gggctggttt gggggggaaca tgtaagccaa gatcaaaaag cagtgagggg
12480
cccgccctga cgactgctgc tcacatctgt gtgtcttgcg caggagacac tgtcggacat
12540
tcgggaaggt gcgccgctag gggtccacct gccttttgtg ggctactcct actcctgcat
12600

ggccctcagg taagcactgc cctggacggc ctccaggggc cacgaggctg cttgagcttc
12660
ctgggtcctg ctccttggca gccaatggag ttgcaggatc agtcttggaa ccttactgtt
12720
ttgggcccaa agactcctaa gaggccagag ttggaggacc ttaaattttc agatctatgt
12780
acttcaaaat gttagattga attttaaaac ctcagagtca cagactgggc ttcccagaat
12840
cttgtaacca ttaactttta cgtctgtagt acacagagcc acaggacttc agaacttgga
12900
aaatatgaag tttagacttt tacaatcagt tgtaaaagaa tgcaaattct ttgaatcagc
12960
catataacaa taaggccatt taaaagtatt aatttaggcg ggccgcggtg gctcacgcct
13020
gtaatcctag cactttggga ggccaaggca ggtggatcat gaggtcagga gatcgagacc
13080
atcctggcta acacggtgaa accccgtctc tactaaaaat acaaaaaaat tagccgggca
13140
tggtggcggg cgcttgcggt cccagctact tgggaggcga ggcaggagaa tggcatgaac
13200
ccgggaggcg gagcttgcag tgagccgaga tcatgccact gcactccagc ctgggcgaca
13260
gagcaagact ccgtctcaaa aaaaaaaaaa aaaaagtatt tatttaggcc gggtgtggtg
13320
gctcacgcct gtaattccag tgctttggga ggatgaggtg ggtggatcac ctgaggtcag
13380
gagttcgaga ccagcctgac caacgtggag aaacctcatc tctactaaaa aacaaaatta
13440
gccaggcgtg gtggcatata cctgtaatcc cagctactca ggaggctgag gcaggagaat
13500
cagaacccag gaggggggagg ttgtggtgag ctgagatcgt gccattgcat tccagcctgg
13560
gcaacaagag tgaaacttca tctcaaaaaa aaaaaaaaaa aagtactaat ttacaggctg
13620
ggcatggtgg ctcacgcttg gaatcccagc actttgggag gctgaagtgg acggattgct
13680
tcagcccagg agttcaagac cagcctgagc aacataatga gaccctgtct ctacaaaaaa
13740
ttgaaaaaat cgtgccaggc atggtggtct gtgcctgcag tcctagctac tcaggagtct
13800
gaagtaggag aatcacttga gcctggagtt tgaggcttca gtgagccatg atagattcca
13860
gcctaggcaa caaagtgaga cctggtctca acaaaagtat taattacaca aataatgcat
13920

tgcttatcac aagtaaatta gaaaatacag ataaggaaaa ggaagttgat atctcgtgag
13980
ctcaccagat ggcagtggtc cctggctcac acgtgtactg acacatgttt aaatagtgga
14040
gaacaggtgt ttttttggtt tgtttttttc cccttcctca tgctactttg tctaagagaa
14100
cagttggttt tctagtcagc ttttattact ggacaacatt acacatacta taccttatca
14160
ttaatgaact ccagcttgat tctgaaccgc tgcggggcct gaacggtggg tcaggattga
14220
acccatcctc tattagaacc caggcgcatg tccaggatag ctaggtcctg agccgtgttc
14280
ccacaggagg gactgctggg ttggagggga cagccacttc ataccccagg gaggagctgt
14340
ccccttccca cagctgagtg gggtgtgctg acctcaagtt gccatcttgg ggtcccatgc
14400
ccagtcttag gaccacatct gtggaggtgg ccagagccaa gcagtctccc catcaggtcg
14460
gcctccctgt cctgaggccc tgagaagagg ggtctgcagc ggtcacatgt caagggagga
14520
gatgagctga ccctagaaca tgggggtctg gaccccaagt ccctgcagaa ggtttagaaa
14580
gagcagctcc caggggccca aggccaggag aggggcaggg cttttcctaa gcagaggagg
14640
ggctattggc ctacctggga ctctgttctc ttcgctctgc tgctcccctt cctcaaatca
14700
ggaggtcttg gaagcagctg cccctaccca caggccagaa gttctggttc tccaccagag
14760
aatcagcatt ctgtctccct ccccactccc tcctcctctc cccagggaca gtgaggtccc
14820
aggccccaca cccatggaac tggaggccga gcagctgctt gagccacacg tgcaagcgcc
14880
cagcctggag ccctcggtgt ccccacagga tgaaacagta agttggtgga ggggaggggg
14940
tccgtcaggg acaattggga gagaaaaggt gagggcttcc cgggtggcgt gcactgtaga
15000
gccctctagg gacttcctga acagaagcag acagaaacca cggagagacg aggttacttc
15060
agacatggga cggtctctgt agttacagtg gggcattaag taagggtgtg tgtgttgctg
15120
gggatctgag aagtcgatct ttgagctgag cgctggtgaa ggagaaacaa gccatggaag
15180
gaaaggtgcc aagtggtcag gcgagagcct ccagggcaaa ggccttgggc aggtgggaat
15240

cctgatttgt tcctgaaagg tagtttggct gaatcattcc tgagaaggct ggagaggcca
15300
gcaggaaaca aaacccagca aggccttttg tcgtgagggc attagggagc tggagggatt
15360
ttgagcagca gagggacata ggttgtgtta gtgtttgagc accagccctc tggtccctgt
15420
gtagatttag aggaccagac tcagggatgg ggctgaggga ggtagggaag ggaggggct
15480
tggatcattg caggagctat ggggattcca gaaatgttga ggggacggag gagtagggga
15540
taaacaagga ttcctagcct ggaaccagtg cccaagtcct gagtcttcca ggagccacag
15600
gcagccttaa gcctggtccc catacacagg ctgaagtggc agttccagcg gctgtccctg
15660
cggcagaggc tgaggccgag gtgacgctgc gggagctcca ggaagccctg gaggaggagg
15720
tgctcacccg gcagagcctg agccgggaga tggaggccat ccgcacggac aaccagaact
15780
tcgccaggtc gggatcgggg ccggggccgg ggccgggatg cgggccggtg gcaacccttg
15840
gcatcccctc tcgtccggcc cggacggact caccgtcctt acctccccac agtcaactac
15900
gcgaggcaga ggctcggaac cgggacctag aggcacacgt ccggcagttg caggagcgga
15960
tggagttgct gcaggcagag ggagccacag gtgagtccct catgtgtccc cttccccgga
16020
ggaccgggag gaggtgggcc gtctgctccg cggggcgtgt atagacacct ggaggaggga
16080
agggacccac gctggggcac gccgcgccac cgccctcctt cgcccctcca cgcgccctat
16140
gcctctttct tctccttcca gctgtcacgg gggtccccag tccccgggcc acggatccac
16200
cttcccatgt aagacccctc tctttccccc gcctcagacc tgctgcccat tctgcagatc
16260
ccctccctgg ctcctggtct ccccgtccag atatagggct caccctacgt ctttgcgact
16320
ttagagggca gaagcccttt attcagcccc agatctccct ccgttcaggc ctcaccagat
16380
tccctccggg atctccctag ataacctccc caacctcgat tcccctcgct gtctctcgcc
16440
ccaccgctga gggctgggct gggctccgat cgggtcacct gtcccttctc tctccagcta
16500
gatggccccc cggccgtggc tgtgggccag tgcccgctgg tggggccagg ccccatgcac
16560

cgccgccacc tgctgctccc tgccagggta cgtccggctg cccacgcccc cctccgccgt
16620
cgcgccccgc gctccacccg ccccttgcca cccgcttagc tgcgcatttg cggggctggg
16680
cccacggcag gagggcggat cttcgggcag ccaatcaaca caggccgcta ggaagcagcc
16740
aatgacgagt tcggacggga ttcgaggcgt gcgagtggac taacaacagc tgtaggctgt
16800
tggggcgggg gcggggcgca gggaagagtg cgggcccacc tatgggcgta ggcggggcga
16860
gtcccaggag ccaatcagag gcccatgccg ggtgttgacc tcgccctctc cccgcaggtc
16920
cctaggcctg gcctatcgga ggcgctttcc ctgctcctgt tcgccgttgt tctgtctcgt
16980
gccgccgccc tgggctgcat tgggttggtg gcccacgccg gccaactcac cgcagtctgg
17040
cgccgcccag gagccgcccg cgctccctga accctagaac tgtcttcgac tccggggccc
17100
cgttggaaga ctgagtgccc ggggcacggc acagaagccg cgcccaccgc ctgccagttc
17160
acaaccgctc cgagcgtggg tctccgccca gctccagtcc tgtgatccgg gcccgccccc
17220
tagcggccgg ggagggaggg gccgggtccg cggccggcga acggggctcg aagggtcctt
17280
gtagccggga atgctgctgc tgctgctgct gctgctgctg ctgctgctgc tgctgctgct
17340
gctgctgctg ctgggggggat cacagaccat ttctttcttt cggccaggct gaggccctga
17400
cgtggatggg caaactgcag gcctgggaag gcagcaagcc gggccgtccg tgttccatcc
17460
tccacgcacc cccacctatc gttggttcgc aaagtgcaaa gctttcttgt gcatgacgcc
17520
ctgctctggg gagcgtctgg cgcgatctct gcctgcttac tcgggaaatt tgcttttgcc
17580
aaacccgctt tttcggggat cccgcgcccc cctcctcact tgcgctgctc tcggagcccc
17640
agccggctcc gcccgcttcg gcggtttgga tatttattga cctcgtcctc cgactcgctg
17700
acaggctaca ggacccccaa caaccccaat ccacgttttg gatgcactga gaccccgaca
17760
ttcctcggta tttattgtct gtccccacct aggacccca cccccgaccc tcgcgaataa
17820
aaggccctcc atctgcccaa agctctggac tccacagtgt ccgcggtttg cgttgtgggc
17880

cggaggctcc gcagcgggcc aatccggagg cgtgtggagg cggccgaagg tctgggagga
17940
gctagcggga tgcgaagcgg ccgaatcagg gttgggggag gaaaagccac ggggcggggc
18000
tttggcgtcc ggccaatagg agggcgagcg ggccacccgg aggcaccgcc cccgcccagc
18060
tgtggcccag ctgtgccacc gagcgtcgag aagagggggc tgggctggca gcgcgcgcgg
18120
ccatcctcct tccactgcgc ctgcgcacgc cacgcgcatc cgctcctggg acgcaagctc
18180
gagaaaagtt gctgcaaact ttctagcccg ttccccgccc ctcctcccgg ccagacccgc
18240
ccccccctgcg gagccgggaa ttccgagggg cggagcgcag gccgagatgg ggaatgtggg
18300
ggcctgcaga ggaccctgga gacggaggcg tgcagaagct cagtctcggg gcggaggctt
18360
cgcgcccttа gtcctcctgg acggcccgtt accttctgcg ttgtcccgat ggggaaactg
18420
aggccctgag ccagaagcac acgctggggg gaggcagaaa gcgcggccag aggcggaggg
18480
aaaacaaagg gagaatcaca gacagacggg agggggacgg acacacacaa ggggacagag
18540
acccgagtgg agagctggat ctcgccttcc cggcgtgggg cgcagggtcg gccagaaaga
18600
agatcgagaa gagcggggag tgggggcgaa aaggggggac aggtggggga ggaggctggg
18660
gaaagcccga gggaggaaga gagggaggga ggaacttccc aaagttgcaa aacatggcta
18720
ccttgcctgc ggagccgagc gcggggccgg cggctggggg ggaggcggtg gcggcggcgg
18780
cggcgaccga agaggaggag gaggaagcgc gccagctctt gcagactttg caggcggccg
18840
agggtgaggc ggcggcggcg gccggggccg gggcgggcgc agcggctgcg ggagctgagg
18900
gcccgggatc cccgggcgtc cccgggtcgc cccccgaggc cgcttccgaa ccgcccacgg
18960
gcctccgctt ctcgcccgag caggtggcgt gcgtctgcga ggcgctgctc caggcgggcc
19020
acgccggccg cttgagccgc ttcctgggcg cactgccccc ggccgagcgc ctacgtggca
19080
gcgacccggt gttgcgcgcg cgggccctgg tggccttcca gcggggcgag tacgccgagc
19140
tctaccggct actcgagagc cgccccttcc ccgccgccca ccacgccttc ctgcaggacc
19200

tctacctgcg cgcgcgctac catgaggccg agcgggcccg cggccgcgcg cttggcgcag
19260
tggacaagta tcgactgcgc aagaagttcc cgctgcccaa gaccatctgg gacggcgagg
19320
agacagtcta ctgcttcaag gagcgctccc gcgcagcgct caaggcctgc taccgcggca
19380
accgctaccc cacgccggac gagaagcgcc gcctggccac actcaccggc ctgtcgctca
19440
cgcaggtcag caactggttc aagaaccggc gacagcgcga ccggaccggg gccggaggcg
19500
gcgcgccctg caagaggtga ggggcctcgg gcggcgcaag tccagctctc ccggggacat
19560
cccgtccacc agccctcttc ccccgtgccc actgctgggg ccggcgcgcc gaggtcctcg
19620
gacatctccc gggaccagct cacaatctca ggcgcccgcg gggcgcgggg actaagtgtg
19680
gacgggacag gcacccgccc gggccctctc cccgcacgcg tctcctcttc cagcggctcc
19740
attccgagct ccttcccaaa tcccatcggt gttggggaat cacactgcgg ggggcactag
19800
agggactgag gaaaaaggac agggcctgtg gccactccac t 19841

<210> 66
<211> 842
<212> PRT
<213> Loài người
<400> 66

Met Ser Arg Pro Leu Ser Asp Gln Glu Lys Arg Lys Gln Ile Ser Val
1 5 10 15

Arg Gly Leu Ala Gly Val Glu Asn Val Thr Glu Leu Lys Lys Asn Phe
20 25 30

Asn Arg His Leu His Phe Thr Leu Val Lys Asp Arg Asn Val Ala Thr
35 40 45

Pro Arg Asp Tyr Tyr Phe Ala Leu Ala His Thr Val Arg Asp His Leu
50 55 60

Val Gly Arg Trp Ile Arg Thr Gln Gln His Tyr Tyr Glu Lys Asp Pro
65 70 75 80

Lys Arg Ile Tyr Tyr Leu Ser Leu Glu Phe Tyr Met Gly Arg Thr Leu

85 90 95

Gln Asn Thr Met Val Asn Leu Ala Leu Glu Asn Ala Cys Asp Glu Ala
100 105 110

Thr Tyr Gln Leu Gly Leu Asp Met Glu Glu Leu Glu Glu Ile Glu Glu
115 120 125

Asp Ala Gly Leu Gly Asn Gly Gly Leu Gly Arg Leu Ala Ala Cys Phe
130 135 140

Leu Asp Ser Met Ala Thr Leu Gly Leu Ala Ala Tyr Gly Tyr Gly Ile
145 150 155 160

Arg Tyr Glu Phe Gly Ile Phe Asn Gln Lys Ile Ser Gly Gly Trp Gln
165 170 175

Met Glu Glu Ala Asp Asp Trp Leu Arg Tyr Gly Asn Pro Trp Glu Lys
180 185 190

Ala Arg Pro Glu Phe Thr Leu Pro Val His Phe Tyr Gly His Val Glu
195 200 205

His Thr Ser Gln Gly Ala Lys Trp Val Asp Thr Gln Val Val Leu Ala
210 215 220

Met Pro Tyr Asp Thr Pro Val Pro Gly Tyr Arg Asn Asn Val Val Asn
225 230 235 240

Thr Met Arg Leu Trp Ser Ala Lys Ala Pro Asn Asp Phe Asn Leu Lys
245 250 255

Asp Phe Asn Val Gly Gly Tyr Ile Gln Ala Val Leu Asp Arg Asn Leu
260 265 270

Ala Glu Asn Ile Ser Arg Val Leu Tyr Pro Asn Asp Asn Phe Phe Glu
275 280 285

Gly Lys Glu Leu Arg Leu Lys Gln Glu Tyr Phe Val Val Ala Ala Thr
290 295 300

Leu Gln Asp Ile Ile Arg Arg Phe Lys Ser Ser Lys Phe Gly Cys Arg
305 310 315 320

Asp Pro Val Arg Thr Asn Phe Asp Ala Phe Pro Asp Lys Val Ala Ile
325 330 335

Gln Leu Asn Asp Thr His Pro Ser Leu Ala Ile Pro Glu Leu Met Arg
340 345 350

Ile Leu Val Asp Leu Glu Arg Met Asp Trp Asp Lys Ala Trp Asp Val
355 360 365

Thr Val Arg Thr Cys Ala Tyr Thr Asn His Thr Val Leu Pro Glu Ala
370 375 380

Leu Glu Arg Trp Pro Val His Leu Leu Glu Thr Leu Leu Pro Arg His
385 390 395 400

Leu Gln Ile Ile Tyr Glu Ile Asn Gln Arg Phe Leu Asn Arg Val Ala
405 410 415

Ala Ala Phe Pro Gly Asp Val Asp Arg Leu Arg Arg Met Ser Leu Val
420 425 430

Glu Glu Gly Ala Val Lys Arg Ile Asn Met Ala His Leu Cys Ile Ala
435 440 445

Gly Ser His Ala Val Asn Gly Val Ala Arg Ile His Ser Glu Ile Leu
450 455 460

Lys Lys Thr Ile Phe Lys Asp Phe Tyr Glu Leu Glu Pro His Lys Phe
465 470 475 480

Gln Asn Lys Thr Asn Gly Ile Thr Pro Arg Arg Trp Leu Val Leu Cys
485 490 495

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

Ile Ser Asp Leu Asp Gln Leu Arg Lys Leu Leu Ser Phe Val Asp Asp
515 520 525

Glu Ala Phe Ile Arg Asp Val Ala Lys Val Lys Gln Glu Asn Lys Leu
530 535 540

Lys Phe Ala Ala Tyr Leu Glu Arg Glu Tyr Lys Val His Ile Asn Pro
545 550 555 560

Asn Ser Leu Phe Asp Ile Gln Val Lys Arg Ile His Glu Tyr Lys Arg
565 570 575

Gln Leu Leu Asn Cys Leu His Val Ile Thr Leu Tyr Asn Arg Ile Lys
580 585 590

Arg Glu Pro Asn Lys Phe Phe Val Pro Arg Thr Val Met Ile Gly Gly
595 600 605

Lys Ala Ala Pro Gly Tyr His Met Ala Lys Met Ile Ile Arg Leu Val
610 615 620

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

Leu Arg Val Ile Phe Leu Glu Asn Tyr Arg Val Ser Leu Ala Glu Lys
645 650 655

Val Ile Pro Ala Ala Asp Leu Ser Glu Gln Ile Ser Thr Ala Gly Thr
660 665 670

Glu Ala Ser Gly Thr Gly Asn Met Lys Phe Met Leu Asn Gly Ala Leu
675 680 685

Thr Ile Gly Thr Met Asp Gly Ala Asn Val Glu Met Ala Glu Glu Ala
690 695 700

Gly Glu Glu Asn Phe Phe Ile Phe Gly Met Arg Val Glu Asp Val Asp
705 710 715 720

Lys Leu Asp Gln Arg Gly Tyr Asn Ala Gln Glu Tyr Tyr Asp Arg Ile
725 730 735

Pro Glu Leu Arg Gln Val Ile Glu Gln Leu Ser Ser Gly Phe Phe Ser

740 745 750

Pro Lys Gln Pro Asp Leu Phe Lys Asp Ile Val Asn Met Leu Met His
755 760 765

His Asp Arg Phe Lys Val Phe Ala Asp Tyr Glu Asp Tyr Ile Lys Cys
770 775 780

Gln Glu Lys Val Ser Ala Leu Tyr Lys Asn Pro Arg Glu Trp Thr Arg
785 790 795 800

Met Val Ile Arg Asn Ile Ala Thr Ser Gly Lys Phe Ser Ser Asp Arg
805 810 815

Thr Ile Ala Gln Tyr Ala Arg Glu Ile Trp Gly Val Glu Pro Ser Arg
820 825 830

Gln Arg Leu Pro Ala Pro Asp Glu Ala Ile
835 840

<210> 67
<211> 429
<212> PRT
<213> Loài người
<400> 67

Met Gln Leu Arg Asn Pro Glu Leu His Leu Gly Cys Ala Leu Ala Leu
1 5 10 15

Arg Phe Leu Ala Leu Val Ser Trp Asp Ile Pro Gly Ala Arg Ala Leu
20 25 30

Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His Trp Glu
35 40 45

Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser Cys Ile
50 55 60

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

Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys Trp Met

85 90 95

Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro Gln Arg
100 105 110

Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser Lys Gly
115 120 125

Leu Lys Leu Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr Cys Ala Gly
130 135 140

Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr Phe Ala
145 150 155 160

Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys Asp Ser
165 170 175

Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala Leu Asn
180 185 190

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

Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr Cys Asn
210 215 220

His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser Ile Lys
225 230 235 240

Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val Asp Val
245 250 255

Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Val Ile Gly Asn
260 265 270

Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu Trp Ala
275 280 285

Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His Ile Ser
290 295 300

Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Val Ile Ala Ile Asn
305 310 315 320

Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly Asp Asn
325 330 335

Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala Val Ala
340 345 350

Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr Ile Ala
355 360 365

Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe Ile
370 375 380

Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu Trp Thr
385 390 395 400

Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu Leu Gln
405 410 415

Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Leu
420 425

<210> 68
<211> 952
<212> PRT
<213> Loài người
<400> 68

Met Gly Val Arg His Pro Pro Cys Ser His Arg Leu Leu Ala Val Cys
1 5 10 15

Ala Leu Val Ser Leu Ala Thr Ala Ala Leu Leu Gly His Ile Leu Leu
20 25 30

His Asp Phe Leu Leu Val Pro Arg Glu Leu Ser Gly Ser Ser Pro Val
35 40 45

Leu Glu Glu Thr His Pro Ala His Gln Gln Gly Ala Ser Arg Pro Gly
50 55 60

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

Gln Cys Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys
85 90 95

Ala Ile Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro
100 105 110

Ala Lys Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe
115 120 125

Phe Pro Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser
130 135 140

Glu Met Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe
145 150 155 160

Pro Lys Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu
165 170 175

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

Val Pro Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu
195 200 205

Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg
210 215 220

Gln Leu Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe
225 230 235 240

Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr
245 250 255

Ile Thr Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser
260 265 270

Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly
275 280 285

Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly
290 295 300

Gly Ser Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val
305 310 315 320

Val Leu Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile
325 330 335

Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln
340 345 350

Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly
355 360 365

Leu Gly Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr
370 375 380

Arg Gln Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val
385 390 395 400

Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe
405 410 415

Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His
420 425 430

Gln Gly Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser
435 440 445

Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg
450 455 460

Gly Val Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val
465 470 475 480

Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu
485 490 495

Ala Trp Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe

500 505 510

Asp Gly Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly
515 520 525

Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val
530 535 540

Pro Gly Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser
545 550 555 560

Ser His Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly
565 570 575

Leu Thr Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly
580 585 590

Thr Arg Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg
595 600 605

Tyr Ala Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu
610 615 620

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

Leu Val Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu
645 650 655

Leu Cys Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg
660 665 670

Asn His Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser
675 680 685

Glu Pro Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala
690 695 700

Leu Leu Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly
705 710 715 720

Glu Thr Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser
725 730 735

Thr Trp Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile
740 745 750

Thr Pro Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro
755 760 765

Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val Pro Val Glu Ala Leu Gly
770 775 780

Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser
785 790 795 800

Glu Gly Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val
805 810 815

His Leu Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr
820 825 830

Thr Thr Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr
835 840 845

Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser
850 855 860

Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala
865 870 875 880

Arg Asn Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly
885 890 895

Ala Gly Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala
900 905 910

Pro Gln Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr
915 920 925

Ser Pro Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly
930 935 940

Glu Gln Phe Leu Val Ser Trp Cys
945 950

<210> 69
<211> 210
<212> PRT
<213> Loài người
<400> 69

Met Trp Thr Leu Gly Arg Arg Ala Val Ala Gly Leu Leu Ala Ser Pro
1 5 10 15

Ser Pro Ala Gln Ala Gln Thr Leu Thr Arg Val Pro Arg Pro Ala Glu
20 25 30

Leu Ala Pro Leu Cys Gly Arg Arg Gly Leu Arg Thr Asp Ile Asp Ala
35 40 45

Thr Cys Thr Pro Arg Arg Ala Ser Ser Asn Gln Arg Gly Leu Asn Gln
50 55 60

Ile Trp Asn Val Lys Lys Gln Ser Val Tyr Leu Met Asn Leu Arg Lys
65 70 75 80

Ser Gly Thr Leu Gly His Pro Gly Ser Leu Asp Glu Thr Thr Tyr Glu
85 90 95

Arg Leu Ala Glu Glu Thr Leu Asp Ser Leu Ala Glu Phe Phe Glu Asp
100 105 110

Leu Ala Asp Lys Pro Tyr Thr Phe Glu Asp Tyr Asp Val Ser Phe Gly
115 120 125

Ser Gly Val Leu Thr Val Lys Leu Gly Gly Asp Leu Gly Thr Tyr Val
130 135 140

Ile Asn Lys Gln Thr Pro Asn Lys Gln Ile Trp Leu Ser Ser Pro Ser
145 150 155 160

Ser Gly Pro Lys Arg Tyr Asp Trp Thr Gly Lys Asn Trp Val Tyr Ser
165 170 175

His Asp Gly Val Ser Leu His Glu Leu Leu Ala Ala Glu Leu Thr Lys
180 185 190

Ala Leu Lys Thr Lys Leu Asp Leu Ser Ser Leu Ala Tyr Ser Gly Lys
195 200 205

Asp Ala
210