



- (12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11) 
CỤC SỞ HỮU TRÍ TUỆ
1-0047369
(51)^{2020.01} A61K 48/00; C12N 15/86; C07K (13) B
14/755; A61K 31/711
-

- (21) 1-2021-04936 (22) 15/01/2020
(86) PCT/US2020/013722 15/01/2020 (87) WO2020/150375 23/07/2020
(30) 62/793,058 16/01/2019 US
(45) 25/06/2025 447 (43) 27/12/2021 405A
(73) 1. BAXALTA INCORPORATED (US)
1200 Lakeside Drive, Bannockburn, IL 60015, United States of America
2. BAXALTA GMBH (CH)
Zaehlerweg 4, 6300 Zug, Switzerland
(72) ROTTENSTEINER, Hanspeter (AT); SCHEIFLINGER, Friedrich (AT).
(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)
-

- (54) HỢP PHẦN AXIT NUCLEIC CÓ CHỨA POLYNUCLEOTIT YẾU TỐ VIII VÀ
VECTƠ LIỆU PHÁP GEN ĐỘNG VẬT CÓ VÚ CÓ CHỨA NÓ

(21) 1-2021-04936

(57) Sáng chế đề xuất, trong số các khía cạnh khác, polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho biến thể Yếu tố VIII để biểu hiện trong tế bào động vật có vú. Theo một số phương án, sáng chế còn đề xuất vector liệu pháp gen động vật có vú để sử dụng trong điều trị bệnh ưa chảy máu A.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến lĩnh vực sinh học phân tử và tế bào. Cụ thể là, sáng chế đề cập đến vectơ virus mã hóa biến thể FVIII yếu tố đông máu tái tổ hợp có sự biểu hiện tăng để sử dụng trong trị liệu gen đối với bệnh ưa chảy máu A.

Tình trạng kỹ thuật của sáng chế

Sự đông máu tiếp diễn thông qua con đường sinh học phức hợp và động lực của các phản ứng sinh hóa phụ thuộc lẫn nhau, được đề cập đến dưới dạng luồng đông máu. Yếu Tố Đông Máu VIII (FVIII) là thành phần chủ chốt trong luồng này. Yếu tố VIII được tuyển mộ đến vị trí chảy máu, và tạo thành phức hợp Xase với Yếu tố IX đã hoạt hóa (FIXa) và Yếu tố X (FX). Phức hợp Xase hoạt hóa FX, mà đến lượt nó hoạt hóa prothrombin thành thrombin, mà sau đó hoạt hóa các thành phần khác trong luồng đông máu để tạo ra cục máu đông ổn định (được xem xét trong Saenko và đồng tác giả, Trends Cardiovasc. Med., 9:185-92 (1999); Lenting và đồng tác giả, Blood, 92:3983-96 (1998)).

Bệnh ưa chảy máu A là rối loạn chảy máu liên kết trên nhiễm sắc thể X bẩm sinh đặc trưng bởi sự thiếu hụt của hoạt tính Yếu tố VIII. Hoạt tính Yếu tố VIII bị giảm đi ức chế vòng phản hồi tích cực trong luồng đông máu. Điều này gây ra sự đông máu không hoàn toàn, mà biểu lộ dưới dạng các đợt chảy máu với khoảng thời gian tăng lên, thâm tím rộng, chảy máu mũi và miệng tự phát, cứng khớp và đau mãn tính, và có khả năng chảy máu bên trong và thiếu máu trong các trường hợp nặng (Zhang và đồng tác giả, Clinic. Rev. Allerg. Immunol., 37:114-24 (2009)).

Thông thường, bệnh ưa chảy máu A được điều trị bằng liệu pháp thay thế Yếu tố VIII, mà gồm có việc dùng protein Yếu tố VIII (ví dụ như, Yếu tố VIII có nguồn gốc huyết tương hoặc được sản xuất theo cách tái tổ hợp) cho cá thể mắc bệnh ưa chảy máu A. Yếu tố VIII được dùng để phòng ngừa để ngăn chặn hoặc làm giảm tần suất của các đợt chảy máu, khi đáp ứng với đợt chảy máu cấp tính, và/hoặc xung quanh khoảng thời

gian phẫu thuật để quản lý sự chảy máu trong quá trình phẫu thuật. Tuy nhiên, có một vài đặc điểm không mong muốn của liệu pháp thay thế Yếu tố VIII.

Thứ nhất, liệu pháp thay thế Yếu tố VIII được sử dụng để điều trị hoặc quản lý bệnh ưa chảy máu A, nhưng không chữa trị sự thiếu hụt Yếu tố VIII ẩn bên dưới. Vì vậy, các cá thể mắc bệnh ưa chảy máu A đòi hỏi liệu pháp thay thế Yếu tố VIII trong toàn bộ cuộc đời của họ. Việc điều trị liên tục rất tốn kém và đòi hỏi cá thể phải duy trì sự tuân thủ nghiêm ngặt, vì việc bỏ lỡ chỉ một ít liều phòng ngừa có thể gây ra hậu quả nghiêm trọng cho cá thể mắc bệnh ưa chảy máu A nghiêm trọng.

Thứ hai, bởi vì Yếu tố VIII có thời gian bán thải tương đối ngắn in vivo, liệu pháp thay thế Yếu tố VIII phòng ngừa thông thường đòi hỏi việc dùng hai ngày hoặc ba ngày một lần. Điều này gây ra gánh nặng lên cá thể để duy trì sự tuân thủ trong toàn bộ cuộc đời của họ. Mặc dù thuốc Yếu tố VIII "tác động lâu dài" thế hệ thứ ba có thể làm giảm tần suất dùng, liệu pháp thay thế Yếu tố FVIII phòng ngừa với các thuốc này vẫn đòi hỏi việc dùng hàng tháng, hàng tuần, hoặc thường xuyên hơn vĩnh viễn. Ví dụ như, việc điều trị phòng ngừa bằng ELOCTATE™ [Yếu Tố Chống Ưa Chảy Máu (Tái Tổ Hợp), Protein Dung Hợp Fc] đòi hỏi việc dùng từ ba đến năm ngày một lần (ELOCTATE™ Prescribing Information, Biogen Idec Inc., (2015)). Hơn nữa, tác dụng lâu dài của chất sinh học được cải biến hóa học (ví dụ như, polypeptit được peg hóa) chưa được hiểu đầy đủ.

Thứ ba, từ 15% đến 30% của tất cả cá thể tiếp nhận liệu pháp thay thế Yếu tố VIII tạo thành kháng thể chất ức chế kháng-Yếu tố VIII, làm cho liệu pháp này không hiệu quả. Liệu pháp đi vòng Yếu tố VIII (ví dụ như, việc dùng chất cô phức hợp prothrombin có nguồn gốc từ huyết tương hoặc được sản xuất theo cách tái tổ hợp) có thể được sử dụng để điều trị bệnh ưa chảy máu ở các cá thể mà tạo thành kháng thể chất ức chế. Tuy nhiên, liệu pháp đi vòng Yếu tố VIII ít hiệu quả hơn liệu pháp thay thế Yếu tố VIII (Mannucci P.M., J Thromb Haemost., 1(7):1349-55 (2003)) và có thể đi kèm với nguy cơ tăng lên của biến chứng tim mạch (Luu và Ewenstein, Haemophilia, 10 Suppl. 2:10-16 (2004)).

Liệu pháp gen soma mang lại hứa hẹn lớn để điều trị bệnh ưa chảy máu A bởi vì nó chữa trị hoạt tính Yếu tố VIII chức năng biểu hiện dưới mức ẩn bên dưới (ví dụ như, do đột biến sai nghĩa hoặc vô nghĩa), chứ không phải là cung cấp liều một lần của hoạt

tính Yếu tố VIII cho cá thể. Bởi vì sự khác nhau trong cơ chế tác động, khi so với liệu pháp thay thế Yếu tố VIII, việc dùng một lần của vector liệu pháp gen Yếu tố VIII có thể cung cấp cho cá thể với Yếu tố VIII trong một vài năm, làm giảm chi phí điều trị và triệt tiêu nhu cầu đối với sự tuân thủ liên tục của bệnh nhân.

Liệu pháp gen Yếu Tố Đông Máu IX (FIX) đã được sử dụng hiệu quả để điều trị các cá thể mắc bệnh ưa chảy máu B, tình trạng đông máu liên quan đặc trưng bởi hoạt tính Yếu tố IX giảm đi (Manno C.S., và đồng tác giả, *Nat Med.*, 12(3):342-47 (2006)). Tuy nhiên, liệu pháp gen Yếu tố VIII có một vài thách thức độc nhất. Ví dụ như, polypeptit Yếu tố VIII kiểu dại, chiều dài đầy đủ (2351 axit amin; mã số truy cập UniProt P00451) lớn hơn năm lần so với polypeptit Yếu tố IX kiểu dại, chiều dài đầy đủ (461 axit amin; mã số truy cập UniProt P00740). Như vậy, trình tự mã hóa của Yếu tố VIII kiểu dại là 7053 cặp bazơ, mà quá lớn để được đóng gói trong vector liệu pháp gen AAV thông thường. Hơn nữa, sự biểu hiện tái tổ hợp đã được báo cáo của biến thể đã được làm khuyết miền B của Yếu tố VIII (BDD-FVIII) là kém. Vì vậy, một vài nhóm đã cố gắng làm thay đổi việc sử dụng bộ ba mã hóa của cấu trúc BDD-FVIII, với sự thành công hạn chế.

Bản chất kỹ thuật của sáng chế

Theo đó, có nhu cầu về biến thể Yếu tố VIII mà trình tự mã hóa của nó được đóng gói hiệu quả hơn vào, và được phân phối thông qua, vector liệu pháp gen. Cũng có nhu cầu về axit nucleic đã được thay đổi bộ ba mã hóa, tổng hợp mà biểu hiện Yếu tố VIII hiệu quả hơn. Cũng có nhu cầu về axit nucleic đã được thay đổi bộ ba mã hóa mã hóa cho polypeptit Yếu tố VIII có tính chất cuộn gập được cải thiện, sự tiết từ tế bào biểu hiện được cải thiện, hoạt tính tăng lên, và/hoặc thời gian bán thải tuần hoàn in vivo được cải thiện, khi so với Yếu tố VIII kiểu dại hoặc Yếu tố VIII bị làm khuyết miền B kiểu dại. Biến thể Yếu tố VIII và axit nucleic đã được thay đổi bộ ba mã hóa này cho phép sự điều trị được cải thiện của sự thiếu hụt Yếu tố VIII (ví dụ như, bệnh ưa chảy máu A). Sự thiếu hụt nêu trên và các vấn đề khác liên quan đến việc điều trị của sự thiếu hụt Yếu tố VIII (ví dụ như, bệnh ưa chảy máu A) được làm giảm hoặc được triệt tiêu bằng biến thể Yếu tố VIII đã được thay đổi bộ ba mã hóa được bộc lộ.

Theo một khía cạnh, hợp phần axit nucleic (ví dụ như, polynucleotit đã được thay đổi bộ ba mã hóa) mã hóa cho biến thể Yếu tố VIII được mô tả. Theo một số phương

án, hợp phần axit nucleic bao gồm polynucleotit có độ tương đồng trình tự cao với các trình tự CS04 (SEQ ID NO:37) hoặc CS12 (SEQ ID NO:1) mã hóa cho biến thể Yếu tố VIII, như được mô tả trong bản mô tả này. Theo một số phương án, hợp phần axit nucleic được mô tả trong bản mô tả này mang lại sự biểu hiện Yếu tố VIII tăng lên và/hoặc hoạt tính Yếu tố VIII tăng lên trong máu của động vật tương quan với trình tự mã hóa Yếu tố VIII kiểu đại và/hoặc trình tự mã hóa Yếu tố VIII được tối ưu hóa bộ ba mã hóa khác. Theo một số phương án, hợp phần axit nucleic còn cho phép sự sản xuất tăng lên của virion liệu pháp gen dựa trên AAV. Theo một số phương án, hợp phần axit nucleic được mô tả trong bản mô tả này có hàm lượng GC giảm và hoặc bao gồm ít dinucleotit CpG hơn, khi so với các trình tự mã hóa cho Yếu tố VIII kiểu đại. Theo một số phương án, biến thể Yếu tố VIII được mã hóa bởi hợp phần axit nucleic được tiết vào máu hiệu quả hơn, in vivo, và/hoặc có thời gian bán thải tuần hoàn tăng lên trong máu, in vivo, tương quan với Yếu tố VIII kiểu đại và/hoặc các biến thể Yếu tố VIII khác.

Theo một số phương án, hợp phần axit nucleic bao gồm polynucleotit mã hóa cho polypeptit Yếu tố VIII có trình tự axit amin của CS12-FL-AA (SEQ ID NO:2), trong đó polynucleotit có trình tự có độ tương đồng trình tự ít nhất là 95% (ví dụ như, độ tương đồng trình tự ít nhất là 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-FL-NA (SEQ ID NO:1).

Theo một số phương án, hợp phần axit nucleic còn bao gồm polynucleotit vùng khởi động được liên kết có điều khiển với polynucleotit Yếu tố VIII, trong đó polynucleotit vùng khởi động có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với hTTR (SEQ ID NO:6). Theo một số phương án, vùng khởi động được gắn trực tiếp vào polynucleotit Yếu tố VIII, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, hợp phần axit nucleic còn bao gồm thành phần đặc hiệu gan được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, thành phần đặc hiệu gan là thành phần vùng tăng cường có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, hoặc 100%) so với CRM8 (SEQ ID NO:5). Theo một số phương án, hợp phần axit nucleic bao gồm hai thành phần đặc hiệu gan này được liên kết có điều khiển với

polynucleotit Yếu tố VIII. Theo một số phương án, hợp phần axit nucleic bao gồm ba thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, một hoặc nhiều thành phần đặc hiệu gan và vùng khởi động được gắn trực tiếp, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, hợp phần axit nucleic có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vr (SEQ ID NO:3).

Theo một số phương án, hợp phần axit có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vrp (SEQ ID NO:10).

Theo một khía cạnh, vector liệu pháp gen động vật có vú mà bao gồm axit nucleic mà mã hóa cho biến thể Yếu tố VIII được mô tả. Theo một số phương án, axit nucleic mà mã hóa cho Yếu tố VIII bao gồm polynucleotit có độ tương đồng trình tự cao với các trình tự CS04 (SEQ ID NO:37) hoặc CS12 (SEQ ID NO:1) mã hóa cho biến thể Yếu tố VIII, như được mô tả trong bản mô tả này. Theo một số phương án, vector liệu pháp gen động vật có vú được mô tả trong bản mô tả này mang lại sự biểu hiện Yếu tố VIII tăng lên và/hoặc hoạt tính Yếu tố VIII tăng lên trong máu của động vật tương quan với vector liệu pháp gen mà bao gồm polynucleotit biến thể Yếu tố VIII được mã hóa tự nhiên hoặc các polynucleotit biến thể Yếu tố VIII được tối ưu hóa bộ ba mã hóa khác. Theo một số phương án, vector liệu pháp gen động vật có vú được mô tả trong bản mô tả này mã hóa cho protein biến thể Yếu tố VIII mà được tiết vào máu hiệu quả hơn, in vivo, và/hoặc có thời gian bán thải tuần hoàn tăng lên trong máu, in vivo, tương quan với Yếu tố VIII kiểu đại và/hoặc các biến thể Yếu tố VIII khác.

Theo một số phương án, vector liệu pháp gen động vật có vú bao gồm polynucleotit mã hóa cho polypeptit Yếu tố VIII có trình tự axit amin của CS12-FL-AA (SEQ ID NO:2), trong đó polynucleotit có trình tự có độ tương đồng trình tự ít nhất là 95% (ví dụ như, độ tương đồng trình tự ít nhất là 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-FL-NA (SEQ ID NO:1).

Theo một số phương án, vector liệu pháp gen động vật có vú còn bao gồm polynucleotit vùng khởi động được liên kết có điều khiển với polynucleotit Yếu tố VIII, trong đó polynucleotit vùng khởi động có trình tự axit nucleic có độ tương đồng trình tự

ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với hTTR (SEQ ID NO:6). Theo một số phương án, vùng khởi động được gắn trực tiếp vào polynucleotit Yếu tố VIII, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, vector liệu pháp gen động vật có vú còn bao gồm thành phần đặc hiệu gan được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, thành phần đặc hiệu gan là thành phần vùng tăng cường có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, hoặc 100%) so với CRM8 (SEQ ID NO:5). Theo một số phương án, hợp phần axit nucleic bao gồm hai thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, hợp phần axit nucleic bao gồm ba thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, một hoặc nhiều thành phần đặc hiệu gan và vùng khởi động được gắn trực tiếp, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, vector liệu pháp gen động vật có vú có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vr (SEQ ID NO:3).

Theo một số phương án, vector liệu pháp gen động vật có vú được bao gồm trong plasmit có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vrp (SEQ ID NO:10).

Theo một khía cạnh, hạt virus kết hợp adeno (AAV) mà bao gồm axit nucleic mà mã hóa cho biến thể Yếu tố VIII được mô tả. Theo một số phương án, axit nucleic mà mã hóa cho Yếu tố VIII bao gồm polynucleotit có độ tương đồng trình tự cao với các trình tự CS04 (SEQ ID NO:37) hoặc CS12 (SEQ ID NO:1) mã hóa cho biến thể Yếu tố VIII, như được mô tả trong bản mô tả này. Theo một số phương án, hạt AAV được mô tả trong bản mô tả này mang lại sự biểu hiện Yếu tố VIII tăng lên và/hoặc hoạt tính Yếu tố VIII tăng lên trong máu của động vật tương quan với hạt AAV mà bao gồm polynucleotit biến thể Yếu tố VIII được mã hóa tự nhiên hoặc các polynucleotit biến thể Yếu tố VIII được tối ưu hóa bộ ba mã hóa khác. Theo một số phương án, hạt AAV được

mô tả trong bản mô tả này mã hóa cho protein biến thể Yếu tố VIII mà được tiết vào máu hiệu quả hơn, in vivo, và/hoặc có thời gian bán thải tuần hoàn tăng lên trong máu, in vivo, tương quan với Yếu tố VIII kiểu đại và/hoặc các biến thể Yếu tố VIII khác.

Theo một số phương án, axit nucleic chứa trong hạt AAV bao gồm polynucleotit mã hóa cho polypeptit Yếu tố VIII có trình tự axit amin của CS12-FL-AA (SEQ ID NO:2), trong đó polynucleotit có trình tự có độ tương đồng trình tự ít nhất là 95% (ví dụ như, độ tương đồng trình tự ít nhất là 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-FL-NA (SEQ ID NO:1).

Theo một số phương án, axit nucleic chứa trong hạt AAV còn bao gồm polynucleotit vùng khởi động được liên kết có điều khiển với polynucleotit Yếu tố VIII, trong đó polynucleotit vùng khởi động có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với hTTR (SEQ ID NO:6). Theo một số phương án, vùng khởi động được gắn trực tiếp vào polynucleotit Yếu tố VIII, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, axit nucleic chứa trong hạt AAV còn bao gồm thành phần đặc hiệu gan được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, thành phần đặc hiệu gan là thành phần vùng tăng cường có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, hoặc 100%) so với CRM8 (SEQ ID NO:5). Theo một số phương án, hợp phần axit nucleic bao gồm hai thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, hợp phần axit nucleic bao gồm ba thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, một hoặc nhiều thành phần đặc hiệu gan và vùng khởi động được gắn trực tiếp, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, axit nucleic chứa trong hạt AAV có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vr (SEQ ID NO:3).

Theo một số phương án, axit nucleic chứa trong hạt AAV được sản xuất bằng cách sử dụng plasmit có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90%

(ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vrp (SEQ ID NO:10).

Theo một khía cạnh, phương pháp điều trị bệnh ưa chảy máu A bằng việc dùng hợp phần axit nucleic mà mã hóa cho biến thể Yếu tố VIII cho bệnh nhân mắc bệnh ưa chảy máu A được mô tả. Theo một số phương án, hợp phần axit nucleic bao gồm polynucleotit có độ tương đồng trình tự cao với các trình tự CS04 (SEQ ID NO:37) hoặc CS12 (SEQ ID NO:1) mã hóa cho biến thể Yếu tố VIII, như được mô tả trong bản mô tả này. Theo một số phương án, phương pháp điều trị bệnh ưa chảy máu A được mô tả trong bản mô tả này dẫn đến sự tăng lên của sự biểu hiện Yếu tố VIII và/hoặc sự tăng lên của hoạt tính Yếu tố VIII trong máu của bệnh nhân mà lớn hơn sự tăng lên của sự biểu hiện Yếu tố VIII và/hoặc sự tăng lên của hoạt tính Yếu tố VIII trong máu của bệnh nhân được dùng hợp phần axit nucleic mà bao gồm trình tự mã hóa Yếu tố VIII kiểu đại và/hoặc trình tự mã hóa Yếu tố VIII được tối ưu hóa bộ ba mã hóa khác. Theo một số phương án, biến thể Yếu tố VIII được mã hóa bởi hợp phần axit nucleic được tiết vào máu hiệu quả hơn, in vivo, và/hoặc có thời gian bán thải tuần hoàn tăng lên trong máu, in vivo, tương quan với Yếu tố VIII kiểu đại và/hoặc các biến thể Yếu tố VIII khác.

Theo một số phương án, hợp phần axit nucleic được dùng bao gồm polynucleotit mã hóa cho polypeptit Yếu tố VIII có trình tự axit amin của CS12-FL-AA (SEQ ID NO:2), trong đó polynucleotit có trình tự có độ tương đồng trình tự ít nhất là 95% (ví dụ như, độ tương đồng trình tự ít nhất là 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-FL-NA (SEQ ID NO:1).

Theo một số phương án, hợp phần axit nucleic được dùng còn bao gồm polynucleotit vùng khởi động được liên kết có điều khiển với polynucleotit Yếu tố VIII, trong đó polynucleotit vùng khởi động có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với hTTR (SEQ ID NO:6). Theo một số phương án, vùng khởi động được gắn trực tiếp vào polynucleotit Yếu tố VIII, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, hợp phần axit nucleic được dùng còn bao gồm thành phần đặc hiệu gan được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, thành phần đặc hiệu gan là thành phần vùng tăng cường có độ tương đồng

trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, hoặc 100%) so với CRM8 (SEQ ID NO:5). Theo một số phương án, hợp phần axit nucleic bao gồm hai thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, hợp phần axit nucleic bao gồm ba thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, một hoặc nhiều thành phần đặc hiệu gan và vùng khởi động được gắn trực tiếp, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, hợp phần axit nucleic được dùng có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vr (SEQ ID NO:3).

Theo một số phương án, hợp phần axit nucleic được dùng có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vrp (SEQ ID NO:10).

Theo một khía cạnh, phương pháp điều trị bệnh ưa chảy máu A bằng việc dùng vector liệu pháp gen động vật có vú mà bao gồm axit nucleic mã hóa cho biến thể Yếu tố VIII được mô tả. Theo một số phương án, axit nucleic mà mã hóa cho Yếu tố VIII bao gồm polynucleotit có độ tương đồng trình tự cao với các trình tự CS04 (SEQ ID NO:37) hoặc CS12 (SEQ ID NO:1) mã hóa cho biến thể Yếu tố VIII, như được mô tả trong bản mô tả này. Theo một số phương án, phương pháp điều trị bệnh ưa chảy máu A được mô tả trong bản mô tả này dẫn đến sự tăng lên của sự biểu hiện Yếu tố VIII và/hoặc sự tăng lên của hoạt tính Yếu tố VIII trong máu của bệnh nhân mà lớn hơn sự tăng lên của sự biểu hiện Yếu tố VIII và/hoặc sự tăng lên của hoạt tính Yếu tố VIII trong máu của bệnh nhân được dùng vector liệu pháp gen động vật có vú mà bao gồm trình tự mã hóa Yếu tố VIII kiểu dại và/hoặc trình tự mã hóa Yếu tố VIII được tối ưu hóa bộ ba mã hóa khác. Theo một số phương án, biến thể Yếu tố VIII được mã hóa bởi axit nucleic ở trong vector liệu pháp gen động vật có vú được tiết vào máu hiệu quả hơn, in vivo, và/hoặc có thời gian bán thải tuần hoàn tăng lên trong máu, in vivo, tương quan với Yếu tố VIII kiểu dại và/hoặc các biến thể Yếu tố VIII khác.

Theo một số phương án, vector liệu pháp gen động vật có vú được dùng bao gồm polynucleotit mã hóa cho polypeptit Yếu tố VIII có trình tự axit amin của CS12-FL-AA (SEQ ID NO:2), trong đó polynucleotit có trình tự có độ tương đồng trình tự ít nhất là 95% (ví dụ như, độ tương đồng trình tự ít nhất là 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-FL-NA (SEQ ID NO:1).

Theo một số phương án, vector liệu pháp gen động vật có vú được dùng còn bao gồm polynucleotit vùng khởi động được liên kết có điều khiển với polynucleotit Yếu tố VIII, trong đó polynucleotit vùng khởi động có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với hTTR (SEQ ID NO:6). Theo một số phương án, vùng khởi động được gắn trực tiếp vào polynucleotit Yếu tố VIII, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, vector liệu pháp gen động vật có vú được dùng còn bao gồm thành phần đặc hiệu gan được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, thành phần đặc hiệu gan là thành phần vùng tăng cường có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, hoặc 100%) so với CRM8 (SEQ ID NO:5). Theo một số phương án, hợp phần axit nucleic bao gồm hai thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, hợp phần axit nucleic bao gồm ba thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, một hoặc nhiều thành phần đặc hiệu gan và vùng khởi động được gắn trực tiếp, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, vector liệu pháp gen động vật có vú được dùng có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vr (SEQ ID NO:3).

Theo một số phương án, vector liệu pháp gen động vật có vú được dùng được bao gồm trong plasmit có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vrp (SEQ ID NO:10).

Theo một khía cạnh, phương pháp điều trị bệnh ưa chảy máu A bằng việc dùng hạt virus kết hợp adeno (AAV) mà bao gồm axit nucleic mã hóa cho biến thể Yếu tố VIII được mô tả. Theo một số phương án, axit nucleic mà mã hóa cho Yếu tố VIII bao gồm polynucleotit có độ tương đồng trình tự cao với các trình tự CS04 (SEQ ID NO:37) hoặc CS12 (SEQ ID NO:1) mã hóa cho biến thể Yếu tố VIII, như được mô tả trong bản mô tả này. Theo một số phương án, phương pháp điều trị bệnh ưa chảy máu A được mô tả trong bản mô tả này dẫn đến sự tăng lên của sự biểu hiện Yếu tố VIII và/hoặc sự tăng lên của hoạt tính Yếu tố VIII trong máu của bệnh nhân mà lớn hơn sự tăng lên của sự biểu hiện Yếu tố VIII và/hoặc sự tăng lên của hoạt tính Yếu tố VIII trong máu của bệnh nhân được dùng hạt virus kết hợp adeno (AAV) mà bao gồm trình tự mã hóa Yếu tố VIII kiểu dại và/hoặc trình tự mã hóa Yếu tố VIII được tối ưu hóa bộ ba mã hóa khác. Theo một số phương án, biến thể Yếu tố VIII được mã hóa bởi axit nucleic ở trong hạt virus kết hợp adeno (AAV) được tiết vào máu hiệu quả hơn, in vivo, và/hoặc có thời gian bán thải tuần hoàn tăng lên trong máu, in vivo, tương quan với Yếu tố VIII kiểu dại và/hoặc các biến thể Yếu tố VIII khác.

Theo một số phương án, axit nucleic chứa trong hạt AAV được dùng bao gồm polynucleotit mã hóa cho polypeptit Yếu tố VIII có trình tự axit amin của CS12-FL-AA (SEQ ID NO:2), trong đó polynucleotit có trình tự có độ tương đồng trình tự ít nhất là 95% (ví dụ như, độ tương đồng trình tự ít nhất là 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-FL-NA (SEQ ID NO:1).

Theo một số phương án, axit nucleic chứa trong hạt AAV được dùng còn bao gồm polynucleotit vùng khởi động được liên kết có điều khiển với polynucleotit Yếu tố VIII, trong đó polynucleotit vùng khởi động có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với hTTR (SEQ ID NO:6). Theo một số phương án, vùng khởi động được gắn trực tiếp vào polynucleotit Yếu tố VIII, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, axit nucleic chứa trong hạt AAV được dùng còn bao gồm thành phần đặc hiệu gan được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, thành phần đặc hiệu gan là thành phần vùng tăng cường có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%,

95%, 96%, 97%, 98%, hoặc 100%) so với CRM8 (SEQ ID NO:5). Theo một số phương án, hợp phần axit nucleic bao gồm hai thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, hợp phần axit nucleic bao gồm ba thành phần đặc hiệu gan này được liên kết có điều khiển với polynucleotit Yếu tố VIII. Theo một số phương án, một hoặc nhiều thành phần đặc hiệu gan và vùng khởi động được gắn trực tiếp, ví dụ như, như được minh họa trên Hình 11.

Theo một số phương án, axit nucleic chứa trong hạt AAV được dùng có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vr (SEQ ID NO:3).

Theo một số phương án, axit nucleic chứa trong hạt AAV được dùng được sản xuất bằng cách sử dụng plasmit có trình tự axit nucleic có độ tương đồng trình tự ít nhất là 90% (ví dụ như, ít nhất là 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100%) so với CS12-CRM8.2-Vrp (SEQ ID NO:10).

Mô tả vắn tắt các hình vẽ

Hình 1A và Hình 1B cùng thể hiện trình tự nucleotit được tối ưu hóa bộ ba mã hóa CS12 (SEQ ID NO:1) mã hóa cho biến thể Yếu tố VIII theo một số phương án (“CS12-FL-NA” đối với trình tự mã hóa chiều dài đầy đủ).

Hình 2 thể hiện trình tự axit amin biến thể Yếu tố VIII (SEQ ID NO: 2) được mã hóa bởi trình tự nucleotit đã được thay đổi bộ ba mã hóa CS12 theo một số phương án (“CS12-FL-AA” đối với trình tự axit amin chiều dài đầy đủ).

Hình 3A và Hình 3B cùng thể hiện trình tự axit nucleic (SEQ ID NO:3) của vector liệu pháp gen đã được làm giảm nucleotit CS12-CRM8.2-Vr mã hóa cho biến thể Yếu tố VIII theo một số phương án.

Hình 4 thể hiện trình tự axit nucleic của các thành phần di truyền khác nhau hữu dụng đối với vector liệu pháp gen mã hóa cho biến thể Yếu tố VIII, bao gồm 5'-ITR (SEQ ID NO:4), thành phần vùng tăng cường CRM8 (SEQ ID NO:5), vùng khởi động TTR người (SEQ ID NO:6), trình tự Kozak tối thiểu (SEQ ID NO:7), và thành phần poly-adenyl hóa tổng hợp (SEQ ID NO:8), và 3'-ITR (SEQ ID NO:9), theo một số phương án.

Hình 5A, Hình 5B, và Hình 5C cùng thể hiện trình tự axit nucleic (SEQ ID NO:10) của plasmit CS12-CRM8.2-Vrp chứa vector liệu pháp gen mã hóa cho biến thể Yếu tố VIII theo một số phương án.

Hình 6A và Hình 6B cùng thể hiện trình tự axit nucleic (SEQ ID NO:38) của vector liệu pháp gen CS12-CRM8.2-V mã hóa cho biến thể Yếu tố VIII theo một số phương án.

Các Hình 7A và 7B cùng thể hiện trình tự nucleotit được tối ưu hóa bộ ba mã hóa CS04 (SEQ ID NO:37) mã hóa cho biến thể Yếu tố VIII theo một số phương án (“CS12-FL-NA” đối với trình tự mã hóa chiều dài đầy đủ).

Hình 8A và Hình 8B cùng thể hiện các trình tự axit amin và nucleotit đối với peptit glycosyl hóa ví dụ mà được cài xen vào cầu nối được thể miền B của biến thể Yếu tố VIII theo một số phương án. “NG1” hoặc NG1-AA” là mã đối với trình tự axit amin, được thể hiện trong dòng trên. “NG1-NA” là mã đối với trình tự axit nucleic, được thể hiện trong dòng dưới đối với mỗi tập hợp. Hình 8A và Hình 8B bộc lộ các trình tự axit amin dưới dạng các SEQ ID NO 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, và 35, và các trình tự nucleotit dưới dạng các SEQ ID NO 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, và 36 tất cả, lần lượt, theo thứ tự xuất hiện.

Hình 9 thể hiện phân tích thẩm tách western của biến thể Yếu tố VIII, có (vNG4/CS04, vNG5/CS04, và vNG16/CS04) và không có (vCS04 với SQ) peptit glycosyl hóa được thiết kế vào trong cầu nối SQ, được biểu hiện ở tế bào Huh-7.

Hình 10 thể hiện phân tích thẩm tách western của biến thể Yếu tố VIII, có (vNG4/CS04, vNG5/CS04, và vNG16/CS04) và không có (vCS04 với SQ) peptit glycosyl hóa được thiết kế vào trong cầu nối SQ, được biểu hiện ở tế bào HepG2 sau khi nhiễm vector liệu pháp gen AAV8.

Hình 11 minh họa cấu trúc liệu pháp gen Yếu tố VIII ví dụ mã hóa cho biến thể Yếu tố VIII theo một số cách thực hiện.

Hình 12 thể hiện mức độ hoạt tính Yếu tố VIII in vivo trong mô hình chuột nhắt dung chịu hFVIII “dòng E”, và in vitro bằng cách sử dụng tế bào HepG2, sau khi nhiễm vector liệu pháp gen AAV8 mã hóa cho protein biến thể Yếu tố VIII theo một số phương án.

Hình 13 thể hiện phân tích điện di gel agarosa của các vectơ liệu pháp gen vCS04, vX5/NG5/CS120, và vX5/NG5/CS12 mã hóa cho biến thể Yếu tố VIII theo một số phương án.

Hình 14 thể hiện mức độ hoạt tính Yếu tố VIII in vivo trong mô hình chuột nhắt knock-in FVIII F17, và in vitro bằng cách sử dụng tế bào HepG2, dòng tế bào gan người (ATCC #HB-8065), sau khi nhiễm vectơ liệu pháp gen AAV8 mã hóa cho protein biến thể Yếu tố VIII theo một số phương án.

Hình 15 thể hiện sơ đồ minh họa của cấu trúc protein Yếu tố VIII người kiểu dại và kiểu Refacto, cũng như là protein Yếu tố VIII được mã hóa bởi polynucleotit CS12, mà chứa đột biến X5 và peptit glycosyl hóa NG5.

Hình 16 thể hiện trình tự axit amin Yếu tố VIII người kiểu dại (SEQ ID NO:39), theo một số phương án (“FVIII-FL-AA”).

Hình 17 thể hiện các trình tự mã hóa ví dụ (các SEQ ID NO 41-53, lần lượt, theo thứ tự xuất hiện) đối với cầu nối được thể miền B theo một số phương án. BDLO04 (SEQ ID NO:41) là một phần của các trình tự nucleotit đã được thay đổi bộ ba mã hóa CS04 mà mã hóa cho cầu nối được thể miền B.

Hình 18 thể hiện khung plasmit ví dụ (SEQ ID NO:54) để tích hợp hệ gen liệu pháp gen Yếu tố VIII, theo một số phương án.

Hình 19 thể hiện đơn vị sao chép ví dụ (“đơn vị sao chép pMB1” – SEQ ID NO:55) và chỉ thị tính kháng ampicilin (“Bla(ApR)” – SEQ ID NO:56) đối với plasmit hữu dụng để tích hợp hệ gen liệu pháp gen Yếu tố VIII, theo một số phương án.

Hình 20A, Hình 20B, và Hình 20C cùng thể hiện trình tự axit nucleic (SEQ ID NO:57) của plasmit CS12-CRM8.2-Vp chứa vectơ liệu pháp gen mã hóa cho biến thể Yếu tố VIII theo một số phương án.

Mô tả chi tiết sáng chế

I. Mở đầu

Liệu pháp gen dựa trên AAV hứa hẹn lớn để điều trị bệnh ưa chảy máu. Đối với bệnh ưa chảy máu B, dữ liệu lâm sàng thứ nhất ủng hộ rằng hàm lượng FIX bằng khoảng 10% có thể được duy trì ở ít nhất là một số bệnh nhân trong thời gian lâu hơn 1 năm.

Tuy nhiên đối với bệnh ưa chảy máu A, việc đạt được mức độ biểu hiện trị liệu bằng 5-10% với vectơ AAV vẫn thách thức vì các lý do khác nhau. Thứ nhất, trình tự mã hóa Yếu tố VIII quá lớn đối với các vectơ dựa trên AAV thông thường. Thứ hai, cấu trúc Yếu tố VIII được làm khuyết hoặc được cắt cụt miền B được thiết kế trải qua sự biểu hiện kém in vivo, ngay cả khi được tối ưu hóa bộ ba mã hóa. Thứ ba, các cấu trúc biến thể Yếu tố VIII được làm khuyết hoặc được cắt cụt miền B này có thời gian bán thải ngắn in vivo, làm trầm trọng thêm tác dụng của sự biểu hiện kém. Thứ tư, ngay cả khi được biểu hiện, FVIII không được tiết hiệu quả từ tế bào, như các yếu tố đông máu khác, chẳng hạn như Yếu tố IX.

Sáng chế đề cập, một phần, đến sự phát hiện vectơ liệu pháp gen chứa trình tự mã hóa biến thể Yếu tố VIII đã được thay đổi bộ ba mã hóa mà giải quyết các vấn đề này và các vấn đề khác liên quan đến liệu pháp gen Yếu tố VIII. Ví dụ như, theo một số phương án, polynucleotit, polypeptit, và cấu trúc liệu pháp gen biến thể Yếu tố VIII được bộc lộ trong bản mô tả này mang lại sự biểu hiện Yếu tố VIII ngoại sinh được cải thiện trong tế bào động vật có vú. Theo một số phương án, polynucleotit, polypeptit, và cấu trúc liệu pháp gen biến thể Yếu tố VIII được bộc lộ trong bản mô tả này mang lại độ sinh khả dụng được cải thiện (ví dụ như, dẫn đến hoạt tính Yếu tố VIII được cải thiện trong máu của bệnh nhân) in vivo. Theo một số phương án, polynucleotit, polypeptit, và cấu trúc liệu pháp gen biến thể Yếu tố VIII được bộc lộ trong bản mô tả này mang lại thời gian bán thải tuần hoàn được cải thiện đối với Yếu tố VIII ngoại sinh trong máu của bệnh nhân. Như được mô tả trong bản mô tả này, một hoặc nhiều ưu điểm trong số các ưu điểm này được thực hiện bằng cách sử dụng dạng kết hợp bất kỳ của một hoặc nhiều sự cải thiện sau đây với hệ thống liệu pháp gen.

Trong một số cách thực hiện, một hoặc nhiều ưu điểm trong số các ưu điểm này được thực hiện bằng cách thiết kế đột biến X5 vào polypeptit biến thể Yếu tố VIII mà được mã hóa bởi polypeptit Yếu tố VIII đã được thay đổi bộ ba mã hóa, như được mô tả trong bản mô tả này. Có lợi là, sự bao gồm của đột biến X5 trong polypeptit Yếu tố VIII được mã hóa, như được mô tả trong bản mô tả này, tạo ra hiệu lực sinh học được cải thiện của Yếu tố VIII được biểu hiện ngoại sinh, in vivo và in vitro. Ví dụ như, như được mô tả trong Ví dụ 4, sự bao gồm của năm đột biến X5 trong miền A1 của chuỗi nặng của polypeptit Refacto FVIII làm tăng hiệu lực sinh học Yếu tố VIII ngoại sinh in vivo ở chuột nhất dòng E2 lên 3 lần sau khi dùng vectơ liệu pháp gen AAV mã hóa cho

biến thể Yếu tố VIII-X5 Refacto, khi so với con chuột nhất được dùng vector liệu pháp gen giống hệt khác mã hóa cho biến thể Yếu tố VIII Refacto kiểu đại (so sánh vX5/CS24 với vCS04 trên Hình 12). Thống nhất với kết quả in vivo, sự bao gồm của năm đột biến X5 trong miền A1 của chuỗi nặng Yếu tố VIII làm tăng hiệu lực sinh học Yếu tố VIII ngoại sinh in vitro lên 4 lần sau khi nhiễm tế bào HepG2 với vector liệu pháp gen AAV mã hóa cho biến thể Yếu tố VIII-X5 Refacto, khi so với tế bào HepG2 đã nhiễm với vector liệu pháp gen giống hệt khác mã hóa cho biến thể Yếu tố VIII Refacto kiểu đại (so sánh vX5/CS24 với vCS04 trên Hình 12).

Trong một số cách thực hiện, một hoặc nhiều ưu điểm trong số các ưu điểm này được thực hiện bằng cách thiết kế peptit glycosyl hóa NG5 vào cầu nối miền B của polypeptit biến thể Yếu tố VIII mà được mã hóa bởi polypeptit Yếu tố VIII đã được thay đổi bộ ba mã hóa, như được mô tả trong bản mô tả này. Có lợi là, sự bao gồm của peptit glycosyl hóa NG5 trong polypeptit Yếu tố VIII được mã hóa, như được mô tả trong bản mô tả này, tạo ra hiệu lực sinh học được cải thiện của Yếu tố VIII được biểu hiện ngoại sinh, in vivo. Ví dụ như, như được mô tả trong Ví dụ 3, sự bao gồm của peptit glycosyl hóa NG trong cầu nối SQ của polypeptit Refacto FVIII làm tăng hiệu lực sinh học Yếu tố VIII ngoại sinh in vivo ở chuột nhất dòng E2 lên 2 lần sau khi dùng vector liệu pháp gen AAV mã hóa cho biến thể Yếu tố VIII-NG5 Refacto, khi so với con chuột nhất được dùng vector liệu pháp gen giống hệt khác mã hóa cho biến thể Yếu tố VIII Refacto kiểu đại (so sánh vNG5/CS04 với vCS04 trên Hình 12).

Trong một số cách thực hiện, một hoặc nhiều ưu điểm trong số các ưu điểm này được thực hiện bằng cách thiết kế cả đột biến X5 và peptit glycosyl hóa NG5 vào polypeptit biến thể Yếu tố VIII mà được mã hóa bởi polypeptit Yếu tố VIII đã được thay đổi bộ ba mã hóa, như được mô tả trong bản mô tả này. Có lợi là, sự bao gồm của đột biến X5 và peptit glycosyl hóa NG5 trong polypeptit Yếu tố VIII được mã hóa, như được mô tả trong bản mô tả này, tạo ra hiệu lực sinh học được cải thiện của Yếu tố VIII được biểu hiện ngoại sinh, in vivo và in vitro. Ví dụ như, như được mô tả trong Ví dụ 5, sự bao gồm của năm đột biến X5 trong miền A1 của chuỗi nặng và peptit glycosyl hóa NG5 trong cầu nối SQ của polypeptit Refacto FVIII làm tăng hiệu lực sinh học Yếu tố VIII ngoại sinh in vivo ở chuột nhất dòng E2 lên 4,5 lần sau khi dùng vector liệu pháp gen AAV mã hóa cho biến thể Yếu tố VIII-X5/NG5 Refacto, khi so với con chuột nhất được dùng vector liệu pháp gen giống hệt khác mã hóa cho biến thể Yếu tố VIII Refacto

kiểu đại (so sánh vX5/NG5/CS125 với vCS04 trên Hình 12). Thống nhất với kết quả in vivo, sự bao gồm của năm đột biến X5 trong miền A1 của chuỗi nặng và peptit glycosyl hóa NG5 trong cầu nối SQ của polypeptit Refacto FVIII làm tăng hiệu lực sinh học Yếu tố VIII ngoại sinh in vitro bằng 3 lần sau khi nhiễm tế bào HepG2 với vector liệu pháp gen AAV mã hóa cho biến thể Yếu tố VIII-X5/NG5 Refacto, khi so với tế bào HepG2 đã nhiễm với vector liệu pháp gen giống hệt khác mã hóa cho biến thể Yếu tố VIII Refacto kiểu đại (so sánh vX5/NG5/CS125 với vCS04 trên Hình 12).

Trong một số cách thực hiện, một hoặc nhiều ưu điểm trong số các ưu điểm này được thực hiện bằng cách sử dụng vùng khởi động hTTR người và một hoặc nhiều thành phần CRM8 đặc hiệu gan ngược dòng của trình tự polynucleotit mã hóa cho polypeptit biến thể Yếu tố VIII. Có lợi là, việc sử dụng vùng khởi động hTTR và một hoặc nhiều thành phần CRM8 đặc hiệu gan tạo ra hiệu lực sinh học được cải thiện của Yếu tố VIII được biểu hiện ngoại sinh trong tế bào người, in vitro. Ví dụ như, như được mô tả trong Ví dụ 2, việc sử dụng vùng khởi động hTTR và một hoặc hai thành phần CRM8 đặc hiệu gan làm tăng hiệu lực sinh học Yếu tố VIII ngoại sinh in vivo trong tế bào HepG2 lần lượt lên khoảng 2 lần và 4 lần, khi so với việc sử dụng các trình tự vùng khởi động và vùng tăng cường TTR chuột nhất (so sánh vCS115 và vCS116 với vCS04 trên Hình 12).

Trong một số cách thực hiện, một hoặc nhiều ưu điểm trong số các ưu điểm này được thực hiện bằng cách loại bỏ các nucleotit ngoại lai nằm giữa các thành phần khác nhau của vector liệu pháp gen AAV mã hóa cho protein biến thể Yếu tố VIII. Có lợi là, việc loại bỏ các nucleotit ngoại lai giữa các thành phần khác nhau tạo ra hiệu lực sinh học được cải thiện của Yếu tố VIII được biểu hiện ngoại sinh trong tế bào người, in vitro. Ví dụ như, như được mô tả trong Ví dụ 6, việc loại bỏ chỉ 71 nucleotit khỏi vector liệu pháp gen AAV vX5/NG5/CS120 mã hóa cho biến thể Yếu tố VIII-X5/NG5 Refacto làm cải thiện hiệu lực sinh học in vitro của biến thể Yếu tố VIII được biểu hiện thêm 50% (so sánh vX5/NG5/CS12 với vX5/NG5/CS120 trên Hình 14).

Trong một số cách thực hiện, một hoặc nhiều ưu điểm trong số các ưu điểm này được thực hiện bằng cách thiết kế cả đột biến X5 và peptit glycosyl hóa NG5 vào polypeptit biến thể Yếu tố VIII mà được mã hóa bởi polypeptit Yếu tố VIII đã được thay đổi bộ ba mã hóa, và bằng cách sử dụng vùng khởi động hTTR người và một hoặc nhiều

thành phần CRM8 đặc hiệu gan ngược dòng của trình tự polynucleotit mã hóa cho polypeptit biến thể Yếu tố VIII. Có lợi là, bằng cách sử dụng dạng kết hợp này của các cải thiện tạo ra hiệu lực sinh học được cải thiện của Yếu tố VIII được biểu hiện ngoại sinh, in vivo và in vitro. Ví dụ như, như được mô tả trong Ví dụ 6, việc sử dụng dạng kết hợp này của các cải thiện trong vector liệu pháp gen AAV làm tăng hiệu lực sinh học in vivo 14,5 lần, tương quan với việc sử dụng vector liệu pháp gen AAV mà bao gồm polynucleotit có cùng sự thay đổi bộ ba mã hóa đối với Yếu tố VIII Refacto kiểu đại được mã hóa bằng cách sử dụng các trình tự vùng khởi động và vùng tăng cường TTR chuột nhắt (so sánh vX5/NG5/CS120 với vSC04 trên Hình 14). Thống nhất với kết quả in vivo, việc sử dụng dạng kết hợp này của các cải thiện trong vector liệu pháp gen AAV làm tăng hiệu lực sinh học in vitro 17 lần, tương quan với việc sử dụng vector liệu pháp gen AAV mà bao gồm polynucleotit có cùng sự thay đổi bộ ba mã hóa đối với Yếu tố VIII Refacto kiểu đại được mã hóa bằng cách sử dụng các trình tự vùng khởi động và vùng tăng cường TTR chuột nhắt (so sánh vX5/NG5/CS120 với vSC04 trên Hình 14). Như báo cáo trong Bảng 4 của WO 2017/083762 (nội dung của nó được kết hợp ở đây trong bản mô tả này để tham khảo), vector vCS04 tạo ra hiệu lực sinh học FVIII in vivo lớn hơn hơn 70 lần, tương quan với vector liệu pháp gen tương đương mã hóa cho polynucleotit Yếu tố VIII Refacto bằng cách sử dụng trình tự mã hóa kiểu đại. Theo đó, mong đợi rằng vector liệu pháp gen vX5/NG5/CS120 sẽ mang lại sự tăng lên từ 1000 lần đến 1250 lần của hiệu lực sinh học FVIII, tương quan với việc sử dụng trình tự mã hóa Refacto kiểu đại.

Trong một số cách thực hiện, một hoặc nhiều ưu điểm trong số các ưu điểm này được thực hiện bằng cách thiết kế cả đột biến X5 và peptit glycosyl hóa NG5 vào polypeptit biến thể Yếu tố VIII mà được mã hóa bởi polypeptit Yếu tố VIII đã được thay đổi bộ ba mã hóa, bằng cách sử dụng vùng khởi động hTTR người và một hoặc nhiều thành phần CRM8 đặc hiệu gan ngược dòng của trình tự polynucleotit mã hóa cho polypeptit biến thể Yếu tố VIII, và việc loại bỏ các nucleotit ngoại lai nằm giữa các thành phần khác nhau của vector liệu pháp gen AAV. Có lợi là, bằng cách sử dụng dạng kết hợp này của các cải thiện tạo ra hiệu lực sinh học được cải thiện của Yếu tố VIII được biểu hiện ngoại sinh, in vivo và in vitro. Ví dụ như, như được mô tả trong Ví dụ 6, việc sử dụng dạng kết hợp này của các cải thiện trong vector liệu pháp gen AAV làm tăng hiệu lực sinh học in vivo 14 lần, tương quan với việc sử dụng vector liệu pháp gen

AAV mà bao gồm polynucleotit có cùng sự thay đổi bộ ba mã hóa đối với Yếu tố VIII Refacto kiểu đại được mã hóa bằng cách sử dụng các trình tự vùng khởi động và vùng tăng cường TTR chuột nhất (so sánh vX5/NG5/CS12 với vSC04 trên Hình 14). Thông nhất với kết quả in vivo, việc sử dụng dạng kết hợp này của các cải thiện trong vector liệu pháp gen AAV làm tăng hiệu lực sinh học in vitro 24 lần, tương quan với việc sử dụng vector liệu pháp gen AAV mà bao gồm polynucleotit có cùng sự thay đổi bộ ba mã hóa đối với Yếu tố VIII Refacto kiểu đại được mã hóa bằng cách sử dụng các trình tự vùng khởi động và vùng tăng cường TTR chuột nhất (so sánh vX5/NG5/CS12 với vSC04 trên Hình 14). Như báo cáo trong Bảng 4 của WO 2017/083762 (nội dung của nó được kết hợp ở đây trong bản mô tả này để tham khảo), vector vCS04 tạo ra hiệu lực sinh học FVIII in vivo lớn hơn hơn 70 lần, tương quan với vector liệu pháp gen tương đương mã hóa cho polynucleotit Yếu tố VIII Refacto bằng cách sử dụng trình tự mã hóa kiểu đại. Theo đó, mong đợi rằng vector liệu pháp gen vX5/NG5/CS120 sẽ mang lại sự tăng lên từ 1000 lần đến 1750 lần của hiệu lực sinh học FVIII, tương quan với việc sử dụng trình tự mã hóa Refacto kiểu đại.

II. Định nghĩa

Như được sử dụng trong bản mô tả này, các thuật ngữ sau đây có nghĩa như được quy cho chúng trừ khi có chỉ dẫn khác.

Như dùng trong bản mô tả này, các thuật ngữ "Yếu tố VIII" và "FVIII" được sử dụng thay thế lẫn nhau, và dùng để chỉ protein bất kỳ có hoạt tính Yếu tố VIII (ví dụ như, hoạt tính FVIII, thường được đề cập đến dưới dạng FVIIIa) hoặc tiền chất protein (ví dụ như, pro-protein hoặc pre-pro-protein) của protein có hoạt tính đồng yếu tố Yếu tố Ixa trong các điều kiện cụ thể, ví dụ như, như đo được bằng cách sử dụng thử nghiệm hoạt hóa Yếu tố X sinh màu hai bước được mô tả trong Chương 2.7.4 của Dược Điển Châu Âu 9.0. Theo phương án ví dụ, polypeptit Yếu tố VIII dùng để chỉ polypeptit mà có trình tự có độ tương đồng trình tự cao (ví dụ như, ít nhất là 70%, 75%, 80%, 85%, 90%, 95%, 99%, hoặc cao hơn) so với chuỗi nặng và chuỗi nhẹ của polypeptit Yếu tố VIII kiểu đại. Theo một số phương án, miền B của polypeptit Yếu tố VIII được làm khuyết, được cắt cụt, hoặc được thay thế bằng cầu nối polypeptit để làm giảm kích thước của polynucleotit mã hóa cho polypeptit Yếu tố VIII.

Ví dụ không làm giới hạn sáng chế về polypeptit Yếu tố VIII kiểu đại bao gồm pre-pro-Yếu tố VIII người (ví dụ như, các mã số truy cập GenBank AAA52485, CAA25619, AAA58466, AAA52484, AAA52420, AAV85964, BAF82636, BAG36452, CAI41660, CAI41666, CAI41672, CAI43241, CAO03404, EAW72645, AAH22513, AAH64380, AAH98389, AAI11968, AAI11970, hoặc AAB61261), pro-Yếu tố VIII tương ứng, và các biến thể tự nhiên của chúng; pre-pro-Yếu tố VIII lợn (ví dụ như, các mã số truy cập UniProt F1RZ36 hoặc K7GSZ5), pro-Yếu tố VIII tương ứng, và các biến thể tự nhiên của chúng; pre-pro-Yếu tố VIII chuột nhắt (ví dụ như, các mã số truy cập GenBank AAA37385, CAM15581, CAM26492, hoặc EDL29229), pro-Yếu tố VIII tương ứng, và các biến thể tự nhiên của chúng; pre-pro-Yếu tố VIII chuột cống (ví dụ như, mã số truy cập GenBank AAQ21580), pro-Yếu tố VIII tương ứng, và các biến thể tự nhiên của chúng; pre-pro-Yếu tố VIII chuột cống; và chất tương tự Yếu tố VIII động vật có vú khác (ví dụ như, khỉ, khỉ hình người, chuột đồng, chuột lang, v.v.).

Như dùng trong bản mô tả này, polypeptit Yếu tố VIII bao gồm các biến thể tự nhiên và cấu trúc nhân tạo có hoạt tính đồng yếu tố Yếu tố IX. Như dùng trong bản bộc lộ này, Yếu tố VIII bao hàm các biến thể tự nhiên bất kỳ, các trình tự thay thế, chất đồng phân, hoặc protein đột biến mà giữ lại một số hoạt tính đồng yếu tố Yếu tố IX cơ bản (ví dụ như, ít nhất là 5%, 10%, 25%, 50%, 75%, hoặc lớn hơn của hoạt tính kiểu đại tương ứng).

Được bao gồm cụ thể ở trong định nghĩa của “Yếu tố VIII” là biến thể Yếu tố VIII, đôi khi còn được đề cập đến dưới dạng “biến thể FVIII”. Protein FVIII biến thể có ít nhất là một sự cải biến axit amin khi so với FVIII kiểu đại người. Ví dụ về sự thay đổi axit amin Yếu tố VIII (tương quan với FVIII-FL-AA (SEQ ID NO: 19)) tìm thấy trong quần thể người bao gồm, nhưng không giới hạn ở, S19R, R22T, Y24C, Y25C, L26P/R, E30V, W33G, Y35C/H, G41C, R48C/K, K67E/N, L69P, E72K, D75E/V/Y, P83R, G89D/V, G92A/V, A97P, E98K, V99D, D101G/H/V, V104D, K108T, M110V, A111T/V, H113R/Y, L117F/R, G121S, E129V, G130R, E132D, Y133C, D135G/Y, T137A/I, S138R, E141K, D145H, V147D, Y155H, V159A, N163K, G164D/V, P165S, C172W, S176P, S179P, V181E/M, K185T, D186G/N/Y, S189L, L191F, G193R, L195P, C198G, S202N/R, F214V, L217H, A219D/T, V220G, D222V, E223K, G224W, T252I, V253F, N254I, G255V, L261P, P262L, G263S, G266F, C267Y, W274C, H275L, G278R, G280D, E284K, V285G, E291G/K, T294I, F295L, V297A, N299I,

R301C/H/L, A303E/P, I307S, S308L, F312S, T314A/I, A315V, G323E, L326P, L327P/V, C329F, I331V, M339T, E340K, V345A/L, C348R/S/Y, Y365C, R391C/H/P, S392L/P, A394S, W401G, I405F/S, E409G, W412G/R, K427I, L431F/S, R437P/W, I438F, G439D/S/V, Y442C, K444R, Y450D/N, T454I, F455C, G466E, P470L/R/T, G474E/R/V, E475K, G477V, D478N, T479R, F484C, A488G, R490G, Y492C/H, Y492H, I494T, P496R, G498R, R503H, G513S/V, I522Y, K529E, W532G, P540T, T541S, D544N, R546W, R550C/G/H, S553P, S554C/G, V556D, R560T, D561G/H/Y, I567T, P569R, S577F, V578A, D579A/H, N583S, Q584H/K/R, I585R/T, M586V, D588G/Y, L594Q, S596P, N601D/K, R602G, S603I/R, W604C, Y605H/S, N609I, R612C, N631K/S, M633I, S635N, N637D/I/S, Y639C, L644V, L650F, V653A/M, L659P, A663V, Q664P, F677L, M681I, V682F, Y683C/N, T686R, F698L, M699T/V, M701I, G705V, G710W, N713I, R717L/W, G720D/S, M721I/L, A723T, L725Q, V727F, E739K, Y742C, R795G, P947R, V1012L, E1057K, H1066Y, D1260E, K1289Q, Q1336K, N1460K, L1481P, A1610S, I1698T, Y1699C/F, E1701K, Q1705H, R1708C/H, T1714S, R1715G, A1720V, E1723K, D1727V, Y1728C, R1740G, K1751Q, F1762L, R1768H, G1769R, L1771P, L1775F/V, L1777P, G1779E/R, P1780L, I1782R, D1788H, M1791T, A1798P, S1799H, R1800C/G/H, P1801A, Y1802C, S1803Y, F1804S, L1808F, M1842I, P1844S, T1845P, E1848G, A1853T/V, S1858C, K1864E, D1865N/Y, H1867P/R, G1869D/V, G1872E, P1873R, L1875P, V1876L, C1877R/Y, L1882P, R1888I, E1894G, I1901F, E1904D/K, S1907C/R, W1908L, Y1909C, A1939T/V, N1941D/S, G1942A, M1945V, L1951F, R1960L/Q, L1963P, S1965I, M1966I/V, G1967D, S1968R, N1971T, H1973L, G1979V, H1980P/Y, F1982I, R1985Q, L1994P, Y1998C, G2000A, T2004R, M2007I, G2013R, W2015C, R2016P/W, E2018G, G2022D, G2028R, S2030N, V2035A, Y2036C, N2038S, 2040Y, G2045E/V, I2051S, I2056N, A2058P, W2065R, P2067L, A2070V, S2082N, S2088F, D2093G/Y, H2101D, T2105N, Q2106E/P/R, G2107S, R2109C, I2117F/S, Q2119R, F2120C/L, Y2124C, R2135P, S2138Y, T2141N, M2143V, F2145C, N2148S, N2157D, P2162L, R2169C/H, P2172L/Q/R, T2173A/I, H2174D, R2178C/H/L, R2182C/H/P, M2183R/V, L2185S/W, S2192I, C2193G, P2196R, G2198V, E2200D, I2204T, I2209N, A2211P, A2220P, P2224L, R2228G/L/P/Q, L2229F, V2242M, W2248C/S, V2251A/E, M2257V, T2264A, Q2265R, F2279C/I, I2281T, D2286G, W2290L, G2304V, D2307A, P2319L/S, R2323C/G/H/L,

R2326G/L/P/Q, Q2330P, W2332R, I2336F, R2339T, G2344C/D/S, và C2345S/Y. Protein Yếu tố VIII còn bao gồm polypeptit chứa sự cải biến sau dịch mã.

Thông thường, polynucleotit mã hóa cho Yếu tố VIII mã hóa cho polypeptit đơn chuỗi không có hoạt tính (ví dụ như, pre-pro-protein) mà trải qua sự xử lý sau dịch mã để tạo thành protein Yếu tố VIII hoạt tính (ví dụ như, FVIIIa). Ví dụ như, tham chiếu đến Hình 15, pre-pro-protein Yếu tố VIII người kiểu dại trước hết được phân cắt để giải phóng peptit tín hiệu được mã hóa (không được thể hiện), tạo thành pro-protein đơn chuỗi thứ nhất (được thể hiện dưới dạng "FVIII kiểu dại người"). Sau đó pro-protein được phân cắt giữa các miền B và A3 để tạo thành polypeptit thứ nhất mà bao gồm chuỗi nặng Yếu tố VIII (ví dụ như, các miền A1 và A2) và miền B, và polypeptit thứ hai mà bao gồm chuỗi nhẹ Yếu tố VIII (ví dụ như, bao gồm các miền A3, C1, và C3). Polypeptit thứ nhất được phân cắt thêm để loại bỏ miền B, và cũng để tách các miền A1 và A2, mà vẫn kết hợp với chuỗi nhẹ Yếu tố VIII trong protein Yếu tố VIIIa trưởng thành. Để xem xét về quy trình trưởng thành hóa Yếu tố VIII, xem Graw và đồng tác giả, *Nat Rev Genet.*, 6(6):488-501 (2005), nội dung của nó được kết hợp trong bản mô tả này để tham khảo đến toàn bộ nội dung của nó cho mọi mục đích.

Như dùng trong bản mô tả này, các thuật ngữ "chuỗi nặng Yếu tố VIII," hoặc đơn giản là "chuỗi nặng," dùng để chỉ khối kết tập của các miền A1 và A2 của polypeptit Yếu tố VIII. Theo phương án ví dụ, các axit amin 20-759 của hFVIII-FL-AA (SEQ ID NO:39) tạo nên chuỗi nặng Yếu tố VIII.

Như dùng trong bản mô tả này, thuật ngữ "chuỗi nhẹ Yếu tố VIII," hoặc đơn giản là "chuỗi nhẹ," dùng để chỉ khối kết tập của các miền A3, C1, và C2 của polypeptit Yếu tố VIII. Theo phương án ví dụ, các axit amin 1668-2351 của hFVIII-FL-AA (SEQ ID NO:39) tạo nên chuỗi nhẹ Yếu tố VIII. Theo một số phương án, chuỗi nhẹ Yếu tố VIII không bao gồm peptit a3 có tính axit, mà được giải phóng trong quá trình trưởng thành *in vivo*.

Thông thường, chuỗi nặng và chuỗi nhẹ Yếu tố VIII được biểu hiện dưới dạng chuỗi polypeptit đơn lẻ, ví dụ như, cùng với miền B hoặc cầu nối được thể miền B tùy chọn. Tuy nhiên, theo một số phương án, chuỗi nặng Yếu tố VIII và chuỗi nhẹ Yếu tố VIII được biểu hiện dưới dạng các chuỗi polypeptit riêng rẽ (ví dụ như, được đồng biểu hiện), và được tái tạo để tạo thành protein Yếu tố VIII (ví dụ như, *in vivo* hoặc *in vitro*).

Như dùng trong bản mô tả này, các thuật ngữ “cầu nối được thể miền B” và “cầu nối Yếu tố VIII” được sử dụng thay thế lẫn nhau, và dùng để chỉ các phiên bản được cắt bớt của miền B Yếu tố VIII kiểu đại (ví dụ như, các axit amin 760-1667 của hFVIII-FL-AA (SEQ ID NO:39)) hoặc peptit được thiết kế để thay thế miền B của polypeptit Yếu tố VIII. Như dùng trong bản mô tả này, cầu nối Yếu tố VIII nằm giữa đầu tận cùng C của chuỗi nặng Yếu tố VIII và đầu tận cùng N của chuỗi nhẹ Yếu tố VIII trong polypeptit biến thể Yếu tố VIII theo một số phương án. Ví dụ không làm giới hạn sáng chế về cầu nối được thể miền B được bộc lộ trong các Bằng Sáng Chế Mỹ Số 4,868,112, 5,112,950, 5,171,844, 5,543,502, 5,595,886, 5,610,278, 5,789,203, 5,972,885, 6,048,720, 6,060,447, 6,114,148, 6,228,620, 6,316,226, 6,346,513, 6,458,563, 6,924,365, 7,041,635, và 7,943,374; các Công Bố Đơn Sáng Chế Mỹ Số 2013/024960, 2015/0071883, và 2015/0158930; và các Công Bố PCT Số WO 2014/064277 và WO 2014/127215, nội dung bộc lộ của chúng được kết hợp ở đây để tham khảo, trong toàn bộ nội dung của chúng, cho mọi mục đích.

Trừ khi có chỉ dẫn khác trong bản mô tả này, việc đánh số của các axit amin Yếu tố VIII dùng để chỉ các axit amin tương ứng trong trình tự Yếu tố VIII người kiểu đại, chiều dài đầy đủ (hFVIII-FL-AA), được thể hiện dưới dạng SEQ ID NO:39 trên Hình 16. Như vậy, khi tham chiếu đến sự thể axit amin trong protein biến thể Yếu tố VIII được bộc lộ trong bản mô tả này, số của axit amin được chỉ ra dùng để chỉ axit amin tương tự (ví dụ như, tương đương về mặt cấu trúc hoặc chức năng) và/hoặc tương đồng (ví dụ như, được bảo toàn trong tiến hóa trong trình tự axit amin sơ cấp) trong trình tự Yếu tố VIII kiểu đại, chiều dài đầy đủ. Ví dụ như, sự thể axit amin T2105N dùng để chỉ sự thể T thành N tại vị trí 2105 của trình tự Yếu tố VIII người kiểu đại, chiều dài đầy đủ (hFVIII-FL-AA; SEQ ID NO:39), sự thể T thành N tại vị trí 1218 của protein biến thể Yếu tố VIII được mã hóa bởi CS12 (CS12-FL-AA; SEQ ID NO: 2).

Như được mô tả trong bản mô tả này, hệ thống đánh số axit amin Yếu tố VIII phụ thuộc vào việc peptit tín hiệu Yếu tố VIII (ví dụ như, các axit amin 1-19 của trình tự Yếu tố VIII người kiểu đại, chiều dài đầy đủ) có được bao gồm hay không. Trong đó peptit tín hiệu được bao gồm, việc đánh số được đề cập đến dưới dạng “bao gồm peptit tín hiệu” hoặc “SPI”. Trong đó peptit tín hiệu không được bao gồm, việc đánh số được đề cập đến dưới dạng “không bao gồm peptit tín hiệu” hoặc “SPE.” Ví dụ như, F328S là đánh số SPI đối với cùng axit amin dưới dạng F309S, trong đánh số SPE. Trừ khi có

chỉ dẫn khác, tất cả các đánh số axit amin dùng để chỉ các axit amin tương ứng trong trình tự Yếu tố VIII người kiểu đại, chiều dài đầy đủ (hFVIII-FL-AA), được thể hiện dưới dạng SEQ ID NO:39 trên Hình 16.

Như được mô tả trong bản mô tả này, polynucleotit đã được thay đổi bộ ba mã hóa mang lại sự biểu hiện tăng của Yếu tố VIII chuyển gen in vivo (ví dụ như, khi được dùng dưới dạng một phần của vector liệu pháp gen), khi so với mức độ biểu hiện Yếu tố VIII tạo ra bởi cấu trúc Yếu tố VIII được mã hóa tự nhiên (ví dụ như, polynucleotit mã hóa cho cùng cấu trúc Yếu tố VIII bằng cách sử dụng bộ ba mã hóa người kiểu đại). Như dùng trong bản mô tả này, thuật ngữ “sự biểu hiện tăng lên” dùng để chỉ mức độ tăng lên của hoạt tính Yếu tố VIII chuyển gen trong máu của động vật được dùng polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho Yếu tố VIII, khi so với mức độ hoạt tính Yếu tố VIII chuyển gen trong máu của động vật được dùng cấu trúc Yếu tố VIII được mã hóa tự nhiên. Mức độ hoạt tính có thể được đo bằng cách sử dụng hoạt tính Yếu tố VIII bất kỳ đã biết trong lĩnh vực, ví dụ như, thử nghiệm hoạt hóa Yếu tố X sinh màu hai bước được mô tả trong Chương 2.7.4 của Dược Điển Châu Âu 9.0. Thử nghiệm ví dụ để xác định hoạt tính Yếu tố VIII là thử nghiệm Technochrome FVIII (Technoclone, Vienna, Áo).

Theo một số phương án, độ sinh khả dụng tăng dùng để chỉ ít nhất là lớn hơn 25% polypeptit Yếu tố VIII chuyển gen trong máu của động vật được dùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, khi so với hàm lượng của polypeptit Yếu tố VIII chuyển gen trong máu của động vật được dùng polynucleotit Yếu tố VIII được mã hóa tự nhiên. Theo một số phương án, độ sinh khả dụng tăng dùng để chỉ ít nhất là lớn hơn 25% polypeptit Yếu tố VIII chuyển gen trong máu của động vật được dùng vector liệu pháp gen được cải thiện mà bao gồm polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, khi so với hàm lượng của polypeptit Yếu tố VIII chuyển gen trong máu của động vật được dùng vector liệu pháp gen khác mà bao gồm cùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa hoặc polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa khác. Theo một số phương án, sự biểu hiện tăng dùng để chỉ lớn hơn ít nhất là 50%, lớn hơn ít nhất là 75%, lớn hơn ít nhất là 100%, lớn hơn ít nhất là 3 lần, lớn hơn ít nhất là 4 lần, lớn hơn ít nhất là 5 lần, lớn hơn ít nhất là 6 lần, lớn hơn ít nhất là 7 lần, lớn hơn ít nhất là 8 lần, lớn hơn ít nhất là 9 lần, lớn hơn ít nhất là 10 lần, lớn hơn ít nhất là 15 lần, lớn hơn ít nhất là 20 lần, lớn hơn ít nhất là 25 lần, lớn hơn ít nhất là 30 lần, lớn

hơn ít nhất là 40 lần, lớn hơn ít nhất là 50 lần, lớn hơn ít nhất là 60 lần, lớn hơn ít nhất là 70 lần, lớn hơn ít nhất là 80 lần, lớn hơn ít nhất là 90 lần, lớn hơn ít nhất là 100 lần, lớn hơn ít nhất là 125 lần, lớn hơn ít nhất là 150 lần, lớn hơn ít nhất là 175 lần, lớn hơn ít nhất là 200 lần, lớn hơn ít nhất là 225 lần, hoặc lớn hơn ít nhất là 250 lần hoạt tính Yếu tố VIII chuyển gen trong máu của động vật được dùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa.

“Hoạt tính Yếu tố VIII” trong bản mô tả này có nghĩa là khả năng để thúc đẩy sự phân cắt của polypeptit Yếu tố X bằng Yếu tố IXa, ví dụ như, hoạt tính đồng yếu tố Yếu tố IXa, thông qua sự thủy phân của liên kết peptit Arg194-Ile195 trong Yếu tố IX kiểu đại, do đó hoạt hóa Yếu tố X thành Yếu tố Xa. Mức độ hoạt tính có thể được đo bằng cách sử dụng hoạt tính Yếu tố VIII bất kỳ đã biết trong lĩnh vực; các thử nghiệm thích hợp được tóm tắt trong bản mô tả này. Thử nghiệm ví dụ để xác định hoạt tính Yếu tố VIII là thử nghiệm hoạt hóa Yếu tố X sinh màu hai bước được mô tả trong Chương 2.7.4 của Dược Điển Châu Âu 9.0.

Như dùng trong bản mô tả này, thuật ngữ “hiệu lực sinh học” dùng để chỉ lượng của hoạt tính Yếu tố VIII trong máu đối tượng, in vivo, hoặc trong dịch nổi bề mặt môi trường nuôi cấy tế bào, in vitro. Theo một số phương án, hiệu lực sinh học sẽ dùng để chỉ lượng của hoạt tính trong mỗi đơn vị thể tích, chẳng hạn như đơn vị của hoạt tính đồng yếu tố Yếu tố XIa trong mỗi ml máu, in vivo, hoặc trong mỗi ml dịch nổi bề mặt môi trường nuôi cấy tế bào, in vitro. Theo một số phương án, hiệu lực sinh học được biểu diễn dưới dạng số lần tăng lên đối với mức độ thứ nhất, ví dụ như, protein Yếu tố VIII được mã hóa tự nhiên hoặc Yếu tố VIII tự nhiên được tối ưu hóa bộ ba mã hóa (ví dụ như, protein FVIII Refecto ‘kiểu đại’). Theo một số phương án, như dùng trong bản mô tả này, hiệu lực sinh học của Yếu tố VIII được biểu hiện ngoại sinh dùng để chỉ lượng của hoạt tính Yếu tố VIII tạo ra bởi protein Yếu tố VIII tái tổ hợp được biểu hiện từ vector liệu pháp gen. Tức là, lượng của hoạt tính Yếu tố VIII trong máu hoặc dịch nổi bề mặt môi trường nuôi cấy tế bào sau khi tính đến lượng đường cơ sở bất kỳ của hoạt tính Yếu tố VIII tự nhiên. Do đó, sự tăng lên của hiệu lực sinh học có thể đạt được bằng một trong hai, hoặc cả hai, việc làm tăng mức độ biểu hiện của protein Yếu tố VIII ngoại sinh và/hoặc việc làm tăng hoạt tính đặc hiệu của protein Yếu tố VIII ngoại sinh, ví dụ như, bằng bao gồm sự thể axit amin (chẳng hạn như đột biến X5) mà mang lại hoạt tính đặc hiệu lớn hơn.

Theo một số phương án, tiềm năng trị liệu của hợp phần polynucleotit Yếu tố VIII được đánh giá bằng sự tăng lên của hoạt tính Yếu tố VIII trong máu của động vật được dùng polynucleotit Yếu tố VIII, ví dụ như, thay vì, hoặc ngoài, sự biểu hiện Yếu tố VIII và/hoặc độ sinh khả dụng tăng lên. Theo một số phương án, như dùng trong bản mô tả này, hoạt tính Yếu tố VIII tăng lên dùng để chỉ sự tăng lớn hơn của hoạt tính Yếu tố VIII trong máu của động vật được dùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, tương quan với hoạt tính Yếu tố VIII đường cơ sở trong máu của động vật trước khi dùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, khi so với sự tăng lên của hoạt tính Yếu tố VIII trong máu của động vật được dùng polynucleotit Yếu tố VIII được mã hóa tự nhiên, tương quan với hoạt tính Yếu tố VIII đường cơ sở trong máu của động vật trước khi dùng polynucleotit Yếu tố VIII được mã hóa tự nhiên. Theo một số phương án, hoạt tính Yếu tố VIII tăng lên dùng để chỉ ít nhất là sự tăng lên hơn 25% của hoạt tính Yếu tố VIII trong máu của động vật được dùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, tương quan với mức độ đường cơ sở của hoạt tính Yếu tố VIII trong máu của động vật trước khi dùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, khi so với sự tăng lên của mức độ hoạt tính Yếu tố VIII trong máu của động vật được dùng polynucleotit Yếu tố VIII được mã hóa tự nhiên, tương quan với mức độ đường cơ sở của hoạt tính Yếu tố VIII trong động vật trước khi dùng polynucleotit Yếu tố VIII được mã hóa tự nhiên. Theo một số phương án, hoạt tính Yếu tố VIII tăng lên dùng để chỉ ít nhất là sự tăng lên hơn 25% của hoạt tính Yếu tố VIII trong máu của động vật được dùng vector liệu pháp gen được cải thiện mà bao gồm polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, tương quan với mức độ đường cơ sở của hoạt tính Yếu tố VIII trong máu của động vật trước khi dùng vector liệu pháp gen được cải thiện, khi so với sự tăng lên của mức độ của hoạt tính Yếu tố VIII trong máu của động vật được dùng vector liệu pháp gen khác mà bao gồm cùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa hoặc polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa khác, tương quan với, tương quan với mức độ đường cơ sở của hoạt tính Yếu tố VIII trong động vật trước khi dùng vector liệu pháp gen khác. Theo một số phương án, hoạt tính Yếu tố VIII tăng lên dùng để chỉ sự tăng lên lớn hơn ít nhất là 50%, lớn hơn ít nhất là 75%, lớn hơn ít nhất là 100%, lớn hơn ít nhất là 3 lần, lớn hơn ít nhất là 4 lần, lớn hơn ít nhất là 5 lần, lớn hơn ít nhất là 6 lần, lớn hơn ít nhất là 7 lần, lớn hơn ít nhất là 8 lần, lớn hơn ít nhất là 9 lần, lớn hơn ít nhất là 10 lần, lớn hơn ít nhất là 15

lần, lớn hơn ít nhất là 20 lần, lớn hơn ít nhất là 25 lần, lớn hơn ít nhất là 30 lần, lớn hơn ít nhất là 40 lần, lớn hơn ít nhất là 50 lần, lớn hơn ít nhất là 60 lần, lớn hơn ít nhất là 70 lần, lớn hơn ít nhất là 80 lần, lớn hơn ít nhất là 90 lần, lớn hơn ít nhất là 100 lần, lớn hơn ít nhất là 125 lần, lớn hơn ít nhất là 150 lần, lớn hơn ít nhất là 175 lần, lớn hơn ít nhất là 200 lần, lớn hơn ít nhất là 225 lần, hoặc lớn hơn ít nhất là 250 lần của hoạt tính Yếu tố VIII trong máu của động vật được dùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, tương quan với mức độ đường cơ sở của hoạt tính Yếu tố VIII trong máu của động vật trước khi dùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, khi so với sự tăng lên của mức độ hoạt tính Yếu tố VIII trong máu của động vật được dùng polynucleotit Yếu tố VIII được mã hóa tự nhiên, hoặc vector liệu pháp gen khác mà bao gồm cùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa hoặc polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa khác, tương quan với mức độ đường cơ sở của hoạt tính Yếu tố VIII trong động vật trước khi dùng polynucleotit Yếu tố VIII được mã hóa tự nhiên hoặc vector liệu pháp gen khác mà bao gồm cùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa hoặc polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa khác. Hoạt tính được đo bằng cách sử dụng thử nghiệm hoạt hóa Yếu tố X sinh màu hai bước được mô tả trong Chương 2.7.4 của Dược Điển Châu Âu 9.0, như được mô tả trong bản mô tả này.

Như được mô tả trong bản mô tả này, polynucleotit đã được thay đổi bộ ba mã hóa mang lại sự sản xuất vector tăng lên, khi so với hàm lượng của sự sản xuất vector tạo ra bởi cấu trúc Yếu tố VIII được mã hóa tự nhiên (ví dụ như, polynucleotit mã hóa cho cùng cấu trúc Yếu tố VIII bằng cách sử dụng bộ ba mã hóa người kiểu dại). Như dùng trong bản mô tả này, thuật ngữ “sự sản xuất virus tăng lên” dùng để chỉ hiệu suất vector tăng lên trong môi trường nuôi cấy tế bào (ví dụ như, độ chuẩn trong mỗi lít môi trường nuôi cấy) được gây nhiễm bằng polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho Yếu tố VIII, khi so với hiệu suất vector trong môi trường nuôi cấy tế bào được gây nhiễm bằng cấu trúc Yếu tố VIII được mã hóa tự nhiên. Hiệu suất vector có thể được đo bằng cách sử dụng thử nghiệm độ chuẩn vector bất kỳ đã biết trong lĩnh vực. Thử nghiệm ví dụ để xác định hiệu suất vector (ví dụ như, của vector AAV) là qPCR nhắm đích đoạn lặp đầu cuối đảo ngược AAV2 (Aurnhammer, Human Gene Therapy Methods: Part B 23:18-28 (2012)).

Theo một số phương án, sự sản xuất virus tăng lên dùng để chỉ hiệu suất vector đã được thay đổi bộ ba mã hóa ít nhất là lớn hơn 25%, khi so với hiệu suất của cấu trúc Yếu tố VIII được mã hóa tự nhiên trong cùng một loại môi trường nuôi cấy. Theo một số phương án, sự sản xuất virus tăng lên dùng để chỉ hiệu suất ít nhất là lớn hơn 25% của vector được cải thiện mà bao gồm polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, khi so với hiệu suất của vector khác mà bao gồm cùng polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa hoặc polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa khác. Theo một số phương án, sự sản xuất vector tăng lên dùng để chỉ hiệu suất vector đã được thay đổi bộ ba mã hóa lớn hơn ít nhất là 50%, lớn hơn ít nhất là 75%, lớn hơn ít nhất là 100%, lớn hơn ít nhất là 3 lần, lớn hơn ít nhất là 4 lần, lớn hơn ít nhất là 5 lần, lớn hơn ít nhất là 6 lần, lớn hơn ít nhất là 7 lần, lớn hơn ít nhất là 8 lần, lớn hơn ít nhất là 9 lần, lớn hơn ít nhất là 10 lần, lớn hơn ít nhất là 15 lần, hoặc lớn hơn ít nhất là 20 lần.

Như dùng trong bản mô tả này, thuật ngữ "bệnh ưa chảy máu" dùng để chỉ nhóm tình trạng bệnh rộng đặc trưng bởi sự đông tụ máu hoặc sự đông máu giảm. Bệnh ưa chảy máu có thể dùng để chỉ bệnh ưa chảy máu Típ A, Típ B, hoặc Típ C, hoặc dùng để chỉ phức hợp của cả ba loại bệnh này. Bệnh ưa chảy máu Típ A (bệnh ưa chảy máu A) gây ra bởi sự giảm đi hoặc sự mất đi của hoạt tính Yếu tố VIII (FVIII) và là dạng nổi bật nhất trong số các kiểu phụ của bệnh ưa chảy máu. Bệnh ưa chảy máu Típ B (bệnh ưa chảy máu B) bắt nguồn từ sự mất đi hoặc sự giảm đi của chức năng đông tụ Yếu tố IX (FIX). Bệnh ưa chảy máu Típ C (bệnh ưa chảy máu C) là hậu quả của sự mất đi hoặc sự giảm đi của hoạt tính đông tụ Yếu tố XI (FXI). Bệnh ưa chảy máu A và B là bệnh liên kết trên nhiễm sắc thể X, mặc dù bệnh ưa chảy máu C là bệnh trên nhiễm sắc thể thường. Việc điều trị thông thường đối với bệnh ưa chảy máu bao gồm cả việc dùng phòng ngừa và theo yêu cầu của yếu tố đông tụ, chẳng hạn như FVIII, FIX, bao gồm Bebulin®-VH, và FXI, cũng như là FEIBA-VH, desmopressin, và dịch truyền huyết tương.

Như dùng trong bản mô tả này, thuật ngữ "liệu pháp gen FVIII" bao gồm phương pháp trị liệu bất kỳ để cung cấp axit nucleic mã hóa cho Yếu tố VIII cho bệnh nhân để làm dịu, làm giảm bớt, hoặc ngăn ngừa sự xuất hiện trở lại của một hoặc nhiều triệu chứng (ví dụ như, yếu tố lâm sàng) liên quan đến bệnh ưa chảy máu. Thuật ngữ này bao hàm việc dùng hợp chất, thuốc, quy trình, hoặc phác đồ bất kỳ có chứa axit nucleic mã

hóa cho phân tử Yếu tố VIII, bao gồm dạng được cải biến bất kỳ của Yếu tố VIII (ví dụ như, biến thể Yếu tố VIII), để duy trì hoặc cải thiện sức khỏe của cá thể mắc bệnh ưa chảy máu. Người có trình độ trung bình trong lĩnh vực sẽ hiểu rằng khoảng thời gian của liệu pháp FVIII hoặc liều lượng của tác nhân trị liệu FVIII có thể được thay đổi, ví dụ như, dựa trên kết quả thu được theo sáng chế.

Như dùng trong bản mô tả này, thuật ngữ “liệu pháp gen Yếu tố VIII,” hoặc “liệu pháp gen FVIII,” bao gồm phương pháp trị liệu bất kỳ cung cấp axit nucleic mã hóa cho polypeptit Yếu tố VIII cho bệnh nhân để làm dịu, làm giảm bớt, hoặc ngăn ngừa sự xuất hiện trở lại của một hoặc nhiều triệu chứng (ví dụ như, yếu tố lâm sàng) liên quan đến sự thiếu hụt Yếu tố VIII (ví dụ như, bệnh ưa chảy máu A). Thuật ngữ này bao hàm việc dùng hợp chất, thuốc, quy trình, hoặc phác đồ bất kỳ có chứa axit nucleic mã hóa cho phân tử Yếu tố VIII, bao gồm dạng được cải biến bất kỳ của Yếu tố VIII (ví dụ như, biến thể Yếu tố VIII có đột biến X5, sự làm khuyết miền B, và/hoặc peptit glycosyl hóa được cài xen vào polypeptit cầu nối miền B), để duy trì hoặc cải thiện sức khỏe của cá thể có sự thiếu hụt Yếu tố VIII (ví dụ như, bệnh ưa chảy máu A). Người có trình độ trung bình trong lĩnh vực sẽ hiểu rằng khoảng thời gian của liệu pháp gen FVIII hoặc liều lượng của liệu pháp gen tác nhân trị liệu FVIII có thể được thay đổi, ví dụ như, dựa trên kết quả thu được theo sáng chế.

Như dùng trong bản mô tả này, thuật ngữ "liệu pháp đi vòng" bao gồm phương pháp trị liệu bất kỳ cung cấp tác nhân cầm máu, hợp chất hoặc yếu tố đông máu không phải là Yếu tố VIII cho bệnh nhân để làm dịu, làm giảm bớt, hoặc ngăn ngừa sự xuất hiện trở lại của một hoặc nhiều triệu chứng (ví dụ như, yếu tố lâm sàng) liên quan đến bệnh ưa chảy máu. Hợp chất và yếu tố đông máu không phải là Yếu tố VIII bao gồm, nhưng không giới hạn ở, Hoạt Tính Đi Vòng Chất Ức Chế Yếu tố VIII (FEIBA), Yếu tố VII đã hoạt hóa tái tổ hợp (FVIIa), chất cô phức hợp prothrombin, và chất cô phức hợp prothrombin đã hoạt hóa. Các hợp chất và yếu tố đông máu không phải là Yếu tố VIII này có thể là tái tổ hợp hoặc có nguồn gốc từ huyết tương. Người có trình độ trung bình trong lĩnh vực sẽ hiểu rằng khoảng thời gian của liệu pháp đi vòng hoặc liều lượng của liệu pháp đi vòng có thể được thay đổi, ví dụ như, dựa trên kết quả thu được theo sáng chế.

Như dùng trong bản mô tả này, “liệu pháp kết hợp” bao gồm việc dùng axit nucleic mã hóa cho phân tử Yếu tố VIII và tác nhân trị liệu bệnh ưa chảy máu A thông thường bao gồm phương pháp trị liệu bất kỳ cung cấp cả axit nucleic mã hóa cho phân tử Yếu tố VIII và phân tử Yếu tố VIII và/hoặc tác nhân cầm máu không phải là Yếu tố VIII (ví dụ như, tác nhân trị liệu đi vòng) cho bệnh nhân để làm dịu, làm giảm bớt, hoặc ngăn ngừa sự xuất hiện trở lại của một hoặc nhiều triệu chứng (ví dụ như, yếu tố lâm sàng) liên quan đến bệnh ưa chảy máu. Thuật ngữ này bao hàm việc dùng hợp chất, thuốc, quy trình, hoặc phác đồ bất kỳ bao gồm axit nucleic mã hóa cho phân tử Yếu tố VIII, bao gồm dạng được cải biến bất kỳ của Yếu tố VIII, mà hữu dụng để duy trì hoặc cải thiện sức khỏe của cá thể mắc bệnh ưa chảy máu và bao gồm tác nhân trị liệu bất kỳ được mô tả trong bản mô tả này.

Các thuật ngữ “lượng hoặc liều lượng hữu hiệu để điều trị” hoặc “liều lượng đủ để điều trị” hoặc “lượng hoặc liều lượng hữu hiệu hoặc đủ” dùng để chỉ liều lượng mà tạo ra tác dụng trị liệu mà nó được dùng cho. Ví dụ như, lượng hữu hiệu để điều trị của thuốc hữu dụng để điều trị bệnh ưa chảy máu có thể là lượng mà có khả năng ngăn ngừa hoặc làm dịu bớt một hoặc nhiều triệu chứng liên quan đến bệnh ưa chảy máu.

Theo một số phương án, việc điều trị hữu hiệu về mặt trị liệu dẫn đến sự giảm đi của tần suất và/hoặc độ nghiêm trọng của việc xảy ra chảy máu ở đối tượng.

Theo một số phương án, việc điều trị hữu hiệu về mặt trị liệu dẫn đến hoạt tính Yếu tố VIII tăng lên trong dòng máu của bệnh nhân, khi so với hoạt tính trước khi điều trị.

Như dùng trong bản mô tả này, thuật ngữ “gen” dùng để chỉ đoạn của phân tử ADN mà mã hóa cho chuỗi polypeptit (ví dụ như, vùng mã hóa). Theo một số phương án, gen được định vị bởi các vùng ngay trước, sau, và/hoặc xen giữa vùng mã hóa mà tham gia vào việc sản xuất chuỗi polypeptit (ví dụ như, các thành phần điều hòa chẳng hạn như vùng khởi động, vùng tăng cường, trình tự polyadenyl hóa, vùng không dịch mã 5’, vùng không dịch mã 3’, hoặc intron).

Như dùng trong bản mô tả này, thuật ngữ “các thành phần điều hòa” dùng để chỉ các trình tự nucleotit, chẳng hạn như vùng khởi động, vùng tăng cường, vùng kết thúc, trình tự polyadenyl hóa, intron, v.v., mà mang lại sự biểu hiện của trình tự mã hóa trong tế bào.

Như dùng trong bản mô tả này, thuật ngữ “thành phần vùng khởi động” dùng để chỉ trình tự nucleotit mà hỗ trợ cho việc điều khiển sự biểu hiện của trình tự mã hóa. Thông thường, thành phần vùng khởi động nằm ở 5’ của vị trí bắt đầu dịch mã của gen. Tuy nhiên, theo các phương án nhất định, thành phần vùng khởi động có thể nằm ở trong trình tự intron, hoặc 3’ của trình tự mã hóa. Theo một số phương án, vùng khởi động hữu dụng đối với vectơ liệu pháp gen có nguồn gốc từ gen tự nhiên của protein đích (ví dụ như, vùng khởi động Yếu tố VIII). Theo một số phương án, vùng khởi động hữu dụng đối với vectơ liệu pháp gen đặc hiệu đối với sự biểu hiện trong tế bào hoặc mô cụ thể của sinh vật đích (ví dụ như, vùng khởi động đặc hiệu gan). Theo phương án khác nữa, một trong số nhiều thành phần vùng khởi động đã được xác định rõ được sử dụng trong vectơ liệu pháp gen được mô tả trong bản mô tả này. Ví dụ không làm giới hạn sáng chế về thành phần vùng khởi động đã được xác định rõ bao gồm vùng khởi động sớm CMV, vùng khởi động β -actin, và vùng khởi động protein liên kết methyl CpG 2 (MeCP2). Theo một số phương án, vùng khởi động là vùng khởi động cấu trúc, mà điều khiển sự biểu hiện về cơ bản không thay đổi của protein đích. Theo phương án khác, vùng khởi động là vùng khởi động cảm ứng, mà điều khiển sự biểu hiện của protein đích khi đáp ứng với kích thích cụ thể (ví dụ như, bộc lộ ra với việc điều trị hoặc tác nhân cụ thể). Để xem xét về thiết kế vùng khởi động đối với liệu pháp gen qua trung gian AAV, xem Gray và đồng tác giả (Human Gene Therapy 22:1143-53 (2011)), nội dung của nó được kết hợp rõ ràng để tham khảo đến toàn bộ nội dung của nó cho mọi mục đích.

Như dùng trong bản mô tả này, thành phần “CRM8” dùng để chỉ phân tử điều hòa tác động cis bắt nguồn từ gen SERPINA1 (mã số truy cập NCBI NM_000295.4) mà tăng cường sự biểu hiện của gen được liên kết có điều khiển, ví dụ như, trình tự mã hóa cho polypeptit Yếu tố VIII, theo kiểu đặc hiệu gan có độ tương đồng trình tự cao với SEQ ID NO:5. Theo một số phương án, thành phần CRM8 giống với SEQ ID NO:5. Như dùng trong bản mô tả này, thành phần CRM8 dùng để chỉ bản sao đơn lẻ của thành phần điều hòa mà, theo một số phương án, được bao gồm trong một hoặc nhiều bản sao ở trong polynucleotit Yếu tố VIII, ví dụ như, 1, 2, 3, hoặc hơn 3 bản sao. Thông tin thêm về thành phần CRM, chẳng hạn như CRM8, xem Chuah MK và đồng tác giả, Mol Ther., 22(9):1605-13 (2014), mà được kết hợp ở đây để tham khảo.

Như dùng trong bản mô tả này, thuật ngữ “được liên kết có điều khiển” dùng để chỉ mối quan hệ giữa trình tự nucleotit tham chiếu thứ nhất (ví dụ như, gen) và trình tự

nucleotit thứ hai (ví dụ như, thành phần điều khiển điều hòa) mà cho phép trình tự nucleotit thứ hai ảnh hưởng đến một hoặc nhiều tính chất liên quan đến trình tự nucleotit tham chiếu thứ nhất (ví dụ như, tốc độ phiên mã). Trong ngữ cảnh của bản bộc lộ này, thành phần điều khiển điều hòa được liên kết có điều khiển với gen chuyển Yếu tố VIII khi thành phần điều hòa được định vị ở trong vector liệu pháp gen sao cho nó gây ra tác dụng (ví dụ như, ảnh hưởng đẩy mạnh hoặc chọn lọc mô) lên sự phiên mã của gen chuyển Yếu tố VIII.

Như dùng trong bản mô tả này, thuật ngữ "vector" dùng để chỉ phương tiện truyền tải bất kỳ dùng để chuyển axit nucleic (ví dụ như, mã hóa cho cấu trúc liệu pháp gen Yếu tố VIII) vào tế bào chủ. Theo một số phương án, vector bao gồm đơn vị sao chép, mà hoạt động để sao chép phương tiện truyền tải, cùng với axit nucleic đích. Ví dụ không làm giới hạn sáng chế về vector hữu dụng đối với liệu pháp gen bao gồm plasmit, phagơ, cosmid, nhiễm sắc thể nhân tạo, và virus, mà hoạt động dưới dạng đơn vị tự sao chép in vivo. Theo một số phương án, vector là phương tiện truyền tải virus để đưa axit nucleic đích (ví dụ như, polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho biến thể Yếu tố VIII) vào. Nhiều virus sinh vật nhân chuẩn đã được cải biến hữu dụng đối với liệu pháp gen đã được biết đến trong lĩnh vực. Ví dụ như, virus kết hợp adeno (AAV) đặc biệt phù hợp để sử dụng trong liệu pháp gen người bởi vì con người là vật chủ tự nhiên đối với virus, virus tự nhiên không được biết là góp phần gây ra bệnh bất kỳ, và virus gây đáp ứng miễn dịch nhẹ.

Như dùng trong bản mô tả này, thuật ngữ “vector virus Yếu tố VIII” dùng để chỉ virus tái tổ hợp có chứa polynucleotit Yếu tố VIII, mã hóa cho polypeptit Yếu tố VIII, mà đủ cho sự biểu hiện của polypeptit Yếu tố VIII khi được đưa vào động vật chủ thích hợp (ví dụ như, con người). Được bao gồm cụ thể ở trong định nghĩa của vector virus Yếu tố VIII là virus tái tổ hợp trong đó polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, mà mã hóa cho polypeptit Yếu tố VIII, đã được cài xen vào trong hệ gen của virus. Cũng được bao gồm cụ thể ở trong định nghĩa của vector virus Yếu tố VIII là virus tái tổ hợp trong đó hệ gen tự nhiên của virus đã được thay thế bằng polynucleotit Yếu tố VIII, mà mã hóa cho polypeptit Yếu tố VIII. Được bao gồm ở trong định nghĩa của vector virus Yếu tố VIII là virus tái tổ hợp có chứa polynucleotit Yếu tố VIII, mà mã hóa cho biến thể của Yếu tố VIII có đột biến X5, sự làm khuyết miễn B, và/hoặc peptit glycosyl hóa được cài xen vào polypeptit cầu nối miễn B.

Như dùng trong bản mô tả này, thuật ngữ “hạt virus Yếu tố VIII” dùng để chỉ hạt virus bao nang polynucleotit Yếu tố VIII, mã hóa cho polypeptit Yếu tố VIII, mà đặc hiệu đối với sự biểu hiện của polypeptit Yếu tố VIII khi được đưa vào động vật chủ thích hợp (ví dụ như, con người). Được bao gồm cụ thể ở trong định nghĩa của hạt virus Yếu tố VIII là hạt virus tái tổ hợp bao nang hệ gen trong đó polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa, mà mã hóa cho polypeptit Yếu tố VIII, đã được cài xen. Cũng được bao gồm cụ thể ở trong định nghĩa của hạt virus Yếu tố VIII là hạt virus tái tổ hợp bao nang polynucleotit Yếu tố VIII, mà mã hóa cho polypeptit Yếu tố VIII, mà thay thế hệ gen tự nhiên của virus. Được bao gồm ở trong định nghĩa của hạt virus Yếu tố VIII là hạt virus tái tổ hợp bao nang polynucleotit Yếu tố VIII, mà mã hóa cho biến thể của Yếu tố VIII có đột biến X5, sự làm khuyết miền B, và/hoặc peptit glycosyl hóa được cài xen vào polypeptit cầu nối miền B.

“AAV” hoặc “virus kết hợp adeno” trong bản mô tả này có nghĩa là Dependoparvovirus ở trong chi Parvoviridae của virus. Như dùng trong bản mô tả này, AAV có thể dùng để chỉ virus bắt nguồn từ hệ gen AAV “kiểu đại” có trong tự nhiên mà polynucleotit Yếu tố VIII đã được cài xen vào đó, virus tái tổ hợp bắt nguồn từ polynucleotit Yếu tố VIII tái tổ hợp được đóng gói vào capsit bằng cách sử dụng protein capsit được mã hóa bởi gen cap AAV có trong tự nhiên, hoặc virus tái tổ hợp bắt nguồn từ polynucleotit Yếu tố VIII tái tổ hợp được đóng gói vào capsit bằng cách sử dụng protein capsit được mã hóa bởi gen cap capsit không tự nhiên. Được bao gồm ở trong định nghĩa của AAV là virus AAV Típ 1 (AAV1), AAV Típ 2 (AAV2), AAV Típ 3 (AAV3), AAV Típ 4 (AAV4), AAV Típ 5 (AAV5), AAV Típ 6 (AAV6), AAV Típ 7 (AAV7), AAV Típ 8 (AAV8), và AAV Típ 9 (AAV9) bao nang polynucleotit Yếu tố VIII và virus được tạo thành bằng một hoặc nhiều protein capsit AAV biến thể bao nang polynucleotit Yếu tố VIII.

“AAV8,” “AAV-8,” hoặc “AAV tít huyết thanh 8” trong bản mô tả này có nghĩa là virus được tạo thành bằng protein virus capsit AAV8 mà bao nang polynucleotit Yếu tố VIII.

Như dùng trong bản mô tả này, thuật ngữ “CpG” dùng để chỉ dinucleotit xytosin-guanin dọc theo sợi đơn của ADN, với “p” thể hiện liên kết photphat giữa hai nucleotit.

Như dùng trong bản mô tả này, thuật ngữ "đảo CpG" dùng để chỉ vùng ở trong polynucleotit có tỉ trọng tăng lên theo thống kê của dinucleotit CpG. Như dùng trong bản mô tả này, vùng của polynucleotit (ví dụ như, polynucleotit mã hóa cho protein Yếu tố VIII đã được thay đổi bộ ba mã hóa) là đảo CpG nếu, trên cửa sổ 200 cặp bazơ: (i) vùng có hàm lượng GC lớn hơn 50%, và (ii) tỉ lệ của dinucleotit CpG được quan sát thấy trong mỗi dinucleotit CpG được mong đợi bằng ít nhất là 0,6, như được xác định bằng mối quan hệ:

$$\frac{N[CpG]*N[chiều dài cửa sổ]}{N[C]*N[G]} \geq 0,6.$$

Đối với thông tin thêm về phương pháp xác định đảo CpG, xem Gardiner-Garden M. và đồng tác giả, J Mol Biol., 196(2):261-82 (1987), nội dung của nó được kết hợp rõ ràng trong bản mô tả này để tham khảo, đến toàn bộ nội dung của nó, cho mọi mục đích.

Như dùng trong bản mô tả này, thuật ngữ "axit nucleic" dùng để chỉ deoxyribonucleotit hoặc ribonucleotit và polyme của chúng ở dạng sợi đơn hoặc sợi kép, và phần bổ sung của chúng. Thuật ngữ này bao hàm axit nucleic chứa dạng tương tự nucleotit đã biết hoặc các gốc hoặc liên kết khung xương được cải biến, mà là dạng tổng hợp, có trong tự nhiên, và không có trong tự nhiên, mà có tính phần liên kết tương tự như axit nucleic tham chiếu, và được chuyển hóa theo cách tương tự với nucleotit tham chiếu. Ví dụ về dạng tương tự này bao gồm, mà không làm giới hạn sáng chế, phosphorothioat, phosphoramidat, methyl phosphonat, chiral-methyl phosphonat, 2-O-methyl ribonucleotit, và peptit-axit nucleic (PNA).

Thuật ngữ "axit amin" dùng để chỉ axit amin có trong tự nhiên, bao gồm axit amin được mã hóa bằng mã di truyền, cũng như là các axit amin mà được cải biến sau đó, ví dụ như, hydroxyprolin, γ -cacboxyglutamat, và O-phosphoserin. Axit amin có trong tự nhiên có thể bao gồm, ví dụ như, D- và L-axit amin. Các axit amin có thể được đề cập đến trong bản mô tả này bằng ký hiệu ba chữ cái đã biết rõ của chúng hoặc bằng ký hiệu một chữ cái theo khuyến nghị của Ủy Ban Danh Pháp Hóa Sinh IUPAC-IUB (IUPAC-IUB Biochemical Nomenclature Commission). Tương tự, các nucleotit có thể được đề cập đến bằng mã một chữ cái được chấp nhận rộng rãi của chúng.

Thuật ngữ "cải biến" trong bản mô tả này nghĩa là sự thế, cài xen, và/hoặc làm khuyết axit amin trong trình tự polypeptit hoặc sự biến đổi đối với gốc được liên kết hóa

học với protein. Ví dụ, cải biến có thể là hydrat cacbon đã được biến đổi hoặc cấu trúc PEG được gắn với protein. Thuật ngữ "cải biến axit amin" trong bản mô tả này nghĩa là sự thế, cài xen, và/hoặc làm khuyết axit amin trong trình tự polypeptit. Để cho rõ ràng, trừ khi có lưu ý khác, cải biến axit amin luôn luôn là cải biến đối với axit amin được mã hóa bởi ADN, ví dụ 20 axit amin có các bộ ba mã hóa trong ADN và ARN.

Thuật ngữ "sự thế axit amin" hoặc "sự thế" trong bản mô tả này nghĩa là sự thay thế axit amin ở vị trí cụ thể trong trình tự polypeptit gốc bằng axit amin khác. Cụ thể, theo một số phương án, sự thế là đối với axit amin không có trong tự nhiên ở vị trí cụ thể, không có trong tự nhiên trong sinh vật này hoặc trong sinh vật bất kỳ. Ví dụ như, sự thế G151K dùng để chỉ polypeptit biến thể, trong đó glyxin tại vị trí 151 được thay thế bằng lysin. Để cho rõ ràng, protein đã được thiết kế di truyền để thay đổi trình tự mã hóa axit nucleic nhưng không thay đổi axit amin ban đầu (ví dụ hoán đổi CGG (mã hóa arginin) với CGA (vẫn mã hóa arginin) để làm tăng mức độ biểu hiện của sinh vật chủ) không phải là "sự thế axit amin"; tức là, mặc dù tạo ra gen mới mã hóa cùng protein, nhưng nếu protein này có cùng axit amin ở vị trí cụ thể mà nó bắt đầu, thì đây không phải là sự thế axit amin. Theo đó, mỗi sự thay đổi của axit nucleic mà mã hóa cho cùng một polypeptit được ngầm hiểu trong mỗi trình tự được mô tả đối với sản phẩm biểu hiện, nhưng không đối với cấu trúc liệu pháp gen thực sự.

Thuật ngữ "sự cài xen axit amin" hoặc "sự cài xen" như được sử dụng trong bản mô tả này nghĩa là việc bổ sung trình tự axit amin ở vị trí cụ thể trong trình tự polypeptit gốc. Ví dụ, -233E hoặc 233E chỉ sự cài xen axit glutamic sau vị trí 233 và trước vị trí 234. Ngoài ra, -233ADE hoặc A233ADE chỉ sự cài xen AlaAspGlu sau vị trí 233 và trước vị trí 234.

Thuật ngữ "sự làm khuyết axit amin" hoặc "sự làm khuyết" như được sử dụng trong bản mô tả này nghĩa là việc loại bỏ trình tự axit amin ở vị trí cụ thể trong trình tự polypeptit gốc. Ví dụ như, E233- hoặc E233#, E233() hoặc E233del chỉ ra sự làm khuyết của axit glutamic tại vị trí 233. Ngoài ra, EDA233- hoặc EDA233# chỉ sự làm khuyết trình tự GluAspAla bắt đầu ở vị trí 233. Các thuật ngữ "tương đồng" hoặc phần trăm "tương đồng," trong ngữ cảnh của hai hoặc nhiều trình tự axit nucleic hoặc peptit, dùng để chỉ hai hoặc nhiều trình tự hoặc trình tự con mà có cùng hoặc có tỉ lệ phần trăm xác định của các gốc axit amin hoặc nucleotit mà giống nhau (tức là, tương đồng khoảng

60%, tốt hơn là tương đồng 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, hoặc cao hơn trên vùng xác định, khi được so sánh và được sắp xếp thẳng hàng để có sự tương ứng lớn nhất trên cửa sổ so sánh hoặc vùng được chỉ định) như đo được bằng cách sử dụng các thuật toán so sánh trình tự BLAST hoặc BLAST 2.0 với các thông số mặc định được mô tả dưới đây, hoặc bằng cách sắp xếp thẳng hàng thủ công và quan sát trực quan.

Như đã biết trong lĩnh vực, nhiều chương trình khác nhau có thể được sử dụng để xác định liệu protein (hoặc axit nucleic như được thảo luận dưới đây) có độ tương đồng hoặc độ tương tự trình tự với trình tự đã biết. Độ tương đồng và/hoặc độ tương tự trình tự được xác định bằng cách sử dụng các kỹ thuật tiêu chuẩn đã biết trong lĩnh vực, bao gồm, nhưng không giới hạn ở, thuật toán độ tương đồng trình tự cục bộ của Smith & Waterman, *Adv. Appl. Math.*, 2:482 (1981), bằng thuật toán sắp xếp thẳng hàng độ tương đồng trình tự của Needleman & Wunsch, *J. Mol. Biol.*, 48:443 (1970), bằng phương pháp tra cứu độ tương tự của Pearson & Lipman, *Proc. Natl. Acad. Sci. U.S.A.*, 85:2444 (1988), bằng cách thực hiện vi tính hóa của các thuật toán này (GAP, BESTFIT, FASTA, và TFASTA trong Gói Phần Mềm Di Truyền Học Wisconsin (Wisconsin Genetics Software Package), Genetics Computer Group, 575 Science Drive, Madison, WI), chương trình trình tự Best Fit được mô tả bởi Devereux và đồng tác giả, *Nucl. Acid Res.*, 12:387-395 (1984), tốt hơn là bằng cách sử dụng thiết lập mặc định, hoặc bằng cách kiểm tra. Tốt hơn là, phần trăm tương đồng được tính bằng FastDB dựa trên các thông số sau đây: phạt bắt cặp nhầm bằng 1; phạt chỗ cách quãng bằng 1; phạt kích thước chỗ cách quãng bằng 0,33; và phạt kết hợp bằng 30, "Current Methods in Sequence Comparison and Analysis," *Macromolecule Sequencing and Synthesis, Selected Methods and Applications*, pp 127-149 (1988), Alan R. Liss, Inc, tất cả chúng được kết hợp để tham khảo.

Ví dụ về thuật toán hữu dụng là PILEUP. PILEUP tạo ra sự sắp xếp thẳng hàng nhiều trình tự từ nhóm của các trình tự liên quan bằng cách sử dụng sự sắp xếp thẳng hàng từng cặp, tiến triển. Nó cũng có thể dựng đồ thị hình cây thể hiện mối quan hệ kết cụm dùng để tạo ra sự sắp xếp thẳng hàng. PILEUP sử dụng sự đơn giản hóa của phương pháp sắp xếp thẳng hàng tiến triển của Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987); phương pháp này tương tự với phương pháp được mô tả bởi Higgins & Sharp *CABIOS* 5:151-153 (1989), đều được kết hợp để tham khảo. Các thông số PILEUP hữu

dụng bao gồm khối lượng chỗ cách quãng mặc định là 3,00, khối lượng chiều dài chỗ cách quãng mặc định là 0,10 và chỗ cách quãng kết thúc có trọng số.

Một ví dụ khác về thuật toán hữu dụng là thuật toán BLAST, được mô tả trong: Altschul và đồng tác giả, *J. Mol. Biol.* 215, 403-410, (1990); Altschul và đồng tác giả, *Nucleic Acids Res.* 25:3389-3402 (1997); và Karlin và đồng tác giả, *Proc. Natl. Acad. Sci. U.S.A.* 90:5873-5787 (1993), đều được kết hợp để tham khảo. Chương trình BLAST đặc biệt hữu dụng là chương trình WU-BLAST-2 mà thu được từ Altschul và đồng tác giả, *Methods in Enzymology*, 266:460-480 (1996); <http://blast.wustl.edu/blast/README.html>]. WU-BLAST-2 sử dụng một vài tham số tìm kiếm, hầu hết trong đó được thiết lập thành các giá trị mặc định. Các thông số có thể điều chỉnh được được đặt ra với các giá trị sau đây: khoảng gổ = 1, đoạn gổ = 0,125, ngưỡng từ (T) = 11. Các tham số HSP S và HSP S2 là các giá trị động và được chỉnh chương trình thiết lập tùy thuộc vào thành phần của trình tự cụ thể và thành phần của cơ sở dữ liệu cụ thể mà trình tự được quan tâm đang được tra cứu trên nó; tuy nhiên, các giá trị này có thể được điều chỉnh để tăng độ nhạy.

Một thuật toán hữu dụng khác là BLAST có chỗ cách quãng, theo báo cáo của Altschul và đồng tác giả, *Nucl. Acids Res.*, 25:3389-3402, được kết hợp để tham khảo. Gapped BLAST sử dụng điểm thế BLOSUM-62; thông số T ngưỡng được đặt là 9; phương pháp hai-tác động để kích hoạt sự kéo dài không cách quãng; tính cho chiều dài quãng cách k chi phí bằng $10+k$; Xu đặt bằng 16, và Xg đặt bằng 40 đối với giai đoạn tra cứu cơ sở dữ liệu và bằng 67 đối với giai đoạn đầu ra của các thuật toán. Sự sắp xếp thẳng hàng có chỗ cách quãng được kích hoạt bằng điểm tương ứng với ~22 bit.

Giá trị % độ tương đồng trình tự axit amin được xác định bằng số lượng của gốc tương đồng trùng khớp chia cho tổng số gốc của trình tự "dài hơn" trong vùng được sắp xếp thẳng hàng. Trình tự "dài hơn" là trình tự có nhiều gốc thực sự nhất trong vùng được sắp xếp thẳng hàng (chỗ cách quãng được đưa vào bằng WU-Blast-2 để tối đa hóa điểm sắp xếp thẳng hàng được bỏ qua). Theo cách tương tự, "phần trăm (%) độ tương đồng trình tự axit nucleic" đối với trình tự mã hóa của polypeptit được xác định được xác định dưới dạng tỉ lệ phần trăm của các gốc nucleotit trong trình tự ứng viên mà tương đồng với các gốc nucleotit trong trình tự mã hóa của protein chu trình tế bào. Phương pháp

được ưu tiên sử dụng mô đun BLASTN của WU-BLAST-2 đặt ở thông số mặc định, với khoảng gôi và đoạn gôi lần lượt đặt ở 1 và 0,125.

Sự sắp xếp thẳng hàng có thể bao gồm sự đưa vào của chỗ cách quãng trong các trình tự cần được sắp xếp thẳng hàng. Ngoài ra, đối với các trình tự mà chứa nhiều hoặc ít axit amin hơn so với protein được mã hóa bởi trình tự của Hình 1 (SEQ ID NO:1), cần hiểu rằng trong một phương án, tỉ lệ phần trăm của độ tương đồng trình tự sẽ được xác định dựa trên số lượng của axit amin hoặc nucleotit tương đồng tương quan với tổng số lượng của axit amin hoặc nucleotit. Do đó, ví dụ như, độ tương đồng trình tự của các trình tự ngắn hơn các trình tự mà được thể hiện trên Hình 1 (SEQ ID NO:1), như được thảo luận dưới đây, sẽ được xác định bằng cách sử dụng số lượng của nucleotit trong trình tự ngắn hơn, trong một phương án. Trong phép tính phần trăm tương đồng khối lượng tương đối không được gán cho các biểu thị khác nhau của sự thay đổi trình tự, chẳng hạn như, sự cài xen, sự làm khuyết, sự thế, v.v..

Theo một phương án, chỉ độ tương đồng được tính điểm dương (+1) và tất cả các dạng của sự thay đổi trình tự bao gồm chỗ cách quãng được gán giá trị bằng "0", mà phòng ngừa nhu cầu đối với thang hoặc thông số có trọng số như được mô tả dưới đây đối với phép tính độ tương tự trình tự. Phần trăm độ tương đồng trình tự có thể được tính, ví dụ như, bằng cách chia số lượng của gốc tương đồng trùng khớp cho tổng số gốc của trình tự "ngắn hơn" trong vùng được sắp xếp thẳng hàng và nhân với 100. Trình tự "dài hơn" là trình tự có nhiều gốc thực sự nhất trong vùng được sắp xếp thẳng hàng.

Thuật ngữ "biến thể alen" dùng để chỉ các dạng đa hình của gen tại locut di truyền cụ thể, cũng như là cADN bắt nguồn từ bản phiên mã mARN của gen, và polypeptit được mã hóa bởi chúng. Thuật ngữ "bộ ba mã hóa động vật có vú được ưu tiên" dùng để chỉ tập hợp con của các bộ ba mã hóa từ trong tập hợp của các bộ ba mã hóa mã hóa cho axit amin mà được sử dụng thường xuyên nhất trong protein được biểu hiện ở tế bào động vật có vú như được chọn từ danh sách sau đây: Gly (GGC, GGG); Glu (GAG); Asp (GAC); Val (GTG, GTC); Ala (GCC, GCT); Ser (AGC, TCC); Lys (AAG); Asn (AAC); Met (ATG); Ile (ATC); Thr (ACC); Trp (TGG); Cys (TGC); Tyr (TAT, TAC); Leu (CTG); Phe (TTC); Arg (CGC, AGG, AGA); Gln (CAG); His (CAC); và Pro (CCC).

Như dùng trong bản mô tả này, thuật ngữ đã được thay đổi bộ ba mã hóa dùng để chỉ trình tự polynucleotit mã hóa cho polypeptit (ví dụ như, protein biến thể Yếu tố VIII), trong đó ít nhất là một bộ ba mã hóa của polynucleotit tự nhiên mã hóa cho polypeptit đã được thay đổi để cải thiện tính chất của trình tự polynucleotit. Theo một số phương án, tính chất được cải thiện này thúc đẩy sự phiên mã tăng của mARN mã hóa cho polypeptit, độ ổn định tăng của mARN (ví dụ như, thời gian bán thải mARN được cải thiện), sự dịch mã tăng của polypeptit, và/hoặc sự đóng gói tăng của polynucleotit ở trong vector. Ví dụ không làm giới hạn sáng chế về các thay đổi mà có thể được sử dụng để đạt được tính chất được cải thiện bao gồm làm thay đổi việc sử dụng và/hoặc sự phân bố của các bộ ba mã hóa đối với axit amin cụ thể, điều chỉnh hàm lượng GC toàn thể và/hoặc cục bộ, việc loại bỏ các trình tự giàu AT, việc loại bỏ thành phần trình tự lặp lại, điều chỉnh hàm lượng dinucleotit CpG toàn thể và/hoặc cục bộ, việc loại bỏ các thành phần điều hòa bí ẩn (ví dụ như, các thành phần hộp TATA và hộp CCAAT), việc loại bỏ vị trí phân cắt intron/exon, việc cải thiện các trình tự điều hòa (ví dụ như, sự đưa vào của trình tự liên ứng Kozak), và việc loại bỏ các thành phần trình tự có khả năng tạo thành cấu trúc thứ cấp (ví dụ như, thân-vòng) trong mARN được phiên mã.

Như thảo luận trong bản mô tả này, có các danh pháp khác nhau dùng để chỉ các thành phần của nội dung bộc lộ trong bản mô tả này. "CS-số" (ví dụ như "CS12", "CS04", v.v.) dùng để chỉ polynucleotit được thay đổi bộ ba mã hóa mã hóa cho polypeptit FVIII và/hoặc polypeptit được mã hóa, bao gồm các biến thể. Ví dụ như, CS12-FL dùng để chỉ trình tự polynucleotit CS12 được thay đổi bộ ba mã hóa Chiều Dài Đầy Đủ hoặc trình tự axit amin (đôi khi được đề cập đến trong bản mô tả này dưới dạng "CS12-FL-AA" đối với trình tự axit amin và "CS12-FL-NA" đối với trình tự axit nucleic) được mã hóa bởi trình tự polynucleotit CS12. Tương tự, "CS12-LC" dùng để chỉ trình tự axit nucleic được thay đổi bộ ba mã hóa ("CS12-LC-NA") mã hóa cho chuỗi nhẹ của polypeptit FVIII hoặc trình tự axit amin (đôi khi còn được đề cập đến trong bản mô tả này dưới dạng "CS12-LC-AA") của chuỗi nhẹ FVIII được mã hóa bởi trình tự polynucleotit CS12. Tương tự, CS12-HC, CS12-HC-AA và CS12-HC-NA là giống nhau đối với chuỗi nặng FVIII. Người có trình độ trung bình trong lĩnh vực hiểu rằng, đối với các cấu trúc chẳng hạn như CS04, mà chỉ được thay đổi bộ ba mã hóa (ví dụ như chúng không chứa sự thể axit amin khác khi so với Refacto), các trình tự axit amin này

là tương đồng, vì các trình tự axit amin không được thay đổi bằng sự tối ưu hóa bộ ba mã hóa. Do đó, các cấu trúc trình tự theo sáng chế bao gồm, nhưng không giới hạn ở, CS12-FL-NA, CS12-FL-AA, CS12-LC-NA, CS12-LC-AA, CS12-HC-AA, và CS12-HC-NA.

Danh pháp này cũng áp dụng cho peptit glycosyl hóa như được thể hiện trên Hình 8, sao cho "NG1-AA" dùng để chỉ trình tự axit amin và NG1-NA dùng để chỉ trình tự axit nucleic.

Sáng chế còn bao gồm biến thể Yếu tố VIII mới khác, như được mô tả dưới đây, với danh pháp thích hợp.

Như dùng trong bản mô tả này, thuật ngữ “sự biểu hiện đặc hiệu gan” dùng để chỉ sự biểu hiện in vivo ưu tiên hoặc chủ yếu của gen cụ thể (ví dụ như, gen Yếu tố VIII chuyển gen, đã được thay đổi bộ ba mã hóa) trong mô gan, khi so với trong các mô khác. Theo một số phương án, sự biểu hiện đặc hiệu gan có nghĩa là ít nhất là 50% của tất cả sự biểu hiện của gen cụ thể xảy ra ở trong mô gan của đối tượng. Theo phương án khác, sự biểu hiện đặc hiệu gan có nghĩa là ít nhất là 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99%, hoặc 100% của tất cả sự biểu hiện của gen cụ thể xảy ra ở trong mô gan của đối tượng. Theo đó, thành phần điều hòa đặc hiệu gan là thành phần điều hòa mà gây ra sự biểu hiện đặc hiệu gan của gen trong mô gan.

Như dùng trong bản mô tả này, các thuật ngữ “ít hơn” X và “ít hơn” X% dùng để chỉ khoảng giá trị từ 0 đến X, không bao gồm giá trị X, ví dụ như, từ 0% đến X%, không bao gồm X%. Như dùng trong bản mô tả này, các thuật ngữ được sử dụng thay thế lẫn nhau với khoảng giá trị bắt đầu ở 0 hoặc 0% đến, nhưng không bao gồm, X hoặc X%.

Như dùng trong bản mô tả này, các thuật ngữ “không lớn hơn” X hoặc “không lớn hơn” X% dùng để chỉ khoảng giá trị từ 0 đến X, bao gồm giá trị X, ví dụ như, từ 0% đến X%, bao gồm X%. Như dùng trong bản mô tả này, các thuật ngữ được sử dụng thay thế lẫn nhau với khoảng giá trị bắt đầu ở 0 hoặc 0% đến, và bao gồm, X hoặc X%.

Như dùng trong bản mô tả này, các thuật ngữ “lớn hơn” X hoặc “lớn hơn” X% dùng để chỉ khoảng giá trị từ X đến giới hạn trên, không bao gồm giá trị X, ví dụ như, từ X% đến 100%, không bao gồm X%. Như dùng trong bản mô tả này, các thuật ngữ

được sử dụng thay thế lẫn nhau với khoảng giá trị bắt đầu ở, nhưng không bao gồm, X hoặc X% đến giới hạn trên mà là 100% trong ngữ cảnh của tỉ lệ phần trăm.

Như dùng trong bản mô tả này, các thuật ngữ “ít nhất là” X hoặc “ít nhất là” X% dùng để chỉ khoảng giá trị từ X đến giới hạn trên, bao gồm giá trị X, ví dụ như, từ X% đến 100%, bao gồm X%. Như dùng trong bản mô tả này, các thuật ngữ được sử dụng thay thế lẫn nhau với khoảng giá trị bắt đầu ở, và bao gồm, X hoặc X% đến giới hạn trên mà là 100% trong ngữ cảnh của tỉ lệ phần trăm.

Như dùng trong bản mô tả này, các thuật ngữ “giữa ‘X’ và ‘Y’,” “giữa ‘X’% và ‘Y’%,” “từ ‘X’ đến ‘Y’,” và “từ ‘X’% đến ‘Y’%” dùng để chỉ khoảng giá trị từ X đến Y, bao gồm các giá trị X và Y, ví dụ như, từ X% đến Y%, bao gồm X% và Y%. Như dùng trong bản mô tả này, các thuật ngữ được sử dụng thay thế lẫn nhau với khoảng giá trị bắt đầu ở X hoặc X% đến, và bao gồm, Y hoặc Y%.

III. Polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa

Theo một số phương án, sáng chế đề xuất polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho biến thể Yếu tố VIII. Các polynucleotit đã được thay đổi bộ ba mã hóa này mang lại hiệu lực sinh học (ví dụ như, hoạt tính) Yếu tố VIII được cải thiện đáng kể khi được dùng trong liệu pháp gen dựa trên cấu trúc AAV. Polynucleotit đã được thay đổi bộ ba mã hóa cũng chứng tỏ sự đóng gói AAV-virion được cải thiện, khi so với cấu trúc được tối ưu hóa bộ ba mã hóa thông thường.

Yếu tố VIII kiểu đại được mã hóa với peptit tín hiệu có 19 axit amin (các axit amin 1-19 của SEQ ID NO:39), mà được phân cắt từ polypeptit được mã hóa trước khi hoạt hóa Yếu tố VIII. Như hiểu rõ bởi người có trình độ trung bình trong lĩnh vực peptit tín hiệu Yếu tố VIII có thể được gây đột biến, được thay thế bằng peptit tín hiệu từ các gen khác hoặc các gen Yếu tố VIII từ sinh vật khác, hoặc được loại bỏ hoàn toàn, mà không ảnh hưởng đến trình tự của polypeptit trưởng thành còn lại sau khi peptit tín hiệu được loại bỏ bằng cách xử lý tế bào.

Theo đó, theo một số phương án, polynucleotit đã được thay đổi bộ ba mã hóa (ví dụ như, hợp phần axit nucleic) được đề xuất trong bản mô tả này có trình tự nucleotit có độ tương đồng trình tự cao với các phần của CS12-FL-NA (SEQ ID NO:1) mã hóa cho chuỗi nặng và chuỗi nhẹ Yếu tố VIII, và cầu nối được thể miền B có 14 axit amin,

ngắn (ví dụ như, cầu nối “SQ” chứa vị trí phân cắt furin để làm thuận lợi cho sự trưởng thành của protein FVIIIa hoạt tính in vivo), mà còn bao gồm một hoặc nhiều trong số năm “đột biến X5” (ví dụ như, một, hai, ba, bốn, hoặc tất cả năm trong số các đột biến I105V/A127S/G151K/M166T/L171P (đánh số SPI; (đánh số SPE là I86V/A108S/G132K/M147T/L152P, lần lượt)), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và/hoặc peptit glycosyl hóa ngắn (ví dụ như, NG5; SEQ ID NO:15) được cài xen vào cầu nối được thể miền B (ví dụ như, cầu nối SQ).

Cụ thể là, bộ đột biến X5 được dựa trên thực tế rằng sự thể của các axit amin lộn 82-176 cho các axit amin người tương ứng trong cấu trúc liệu pháp gen được làm khuyết miền B làm tăng hoạt tính Yếu tố VIII khi được biểu hiện ở tế bào HEK293 (W. Xiao, truyền thông). Đột biến ngược của các axit amin lộn đơn lẻ vào cấu trúc BDD-FVIII người xác định năm axit amin ở trong miền A1 mà gây nên hiện tượng này: I105V, A127S, G151K, M166T, và L171P (SPI). Sự đưa vào của dạng kết hợp của các đột biến này vào cấu trúc của người lặp lại hoạt tính được cải thiện của sự thể của lộn lớn hơn. (W. Xiao, truyền thông). Theo đó, theo một số phương án, polypeptit Yếu tố VIII được mã hóa bao gồm một hoặc nhiều sự thể axit amin được chọn từ I105V, A127S, G151K, M166T, và L171P, với toàn bộ bộ 5 axit amin được sử dụng cụ thể trong nhiều phương án.

Polynucleotit được thay đổi bộ ba mã hóa CS12

Theo một số phương án, hợp phần axit nucleic được đề xuất trong bản mô tả này bao gồm polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa mà có trình tự nucleotit có độ tương đồng trình tự cao với CS12-FL-NA (SEQ ID NO:1) và mã hóa cho polypeptit Yếu tố VIII có chuỗi nặng và chuỗi nhẹ Yếu tố VIII người, và cầu nối được thể miền B có 14 axit amin, ngắn (cầu nối “SQ”) chứa vị trí phân cắt furin để làm thuận lợi cho sự trưởng thành của protein FVIIIa hoạt tính in vivo, trong đó chuỗi nặng của polypeptit Yếu tố VIII bao gồm năm đột biến X5 (I105V, A127S, G151K, M166T, và L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ.

Theo một số phương án, polynucleotit Yếu tố VIII có trình tự nucleotit có độ tương đồng ít nhất là 95% với CS12-FL-NA (SEQ ID NO:1). Theo một số phương án, trình tự nucleotit có độ tương đồng ít nhất là 96% với CS12-FL-NA (SEQ ID NO:1).

Theo một số phương án, trình tự nucleotit có độ tương đồng ít nhất là 97% với CS12-FL-NA (SEQ ID NO:1). Theo một số phương án, trình tự nucleotit có độ tương đồng ít nhất là 98% với CS12-FL-NA (SEQ ID NO:1). Theo một số phương án, trình tự nucleotit có độ tương đồng ít nhất là 99% với CS12-FL-NA (SEQ ID NO:1). Theo một số phương án, trình tự nucleotit có độ tương đồng ít nhất là 99,5% với CS12-FL-NA (SEQ ID NO:1). Theo một số phương án, trình tự nucleotit có độ tương đồng ít nhất là 99,9% với CS12-FL-NA (SEQ ID NO:1). Theo một số phương án, trình tự nucleotit giống với CS12-FL-NA (SEQ ID NO:1).

Theo một số phương án, biến thể Yếu tố VIII được mã hóa bởi polynucleotit đã được thay đổi bộ ba mã hóa có trình tự axit amin có độ tương đồng trình tự cao với CS12-FL-AA (SEQ ID NO: 2), bao gồm năm đột biến X5 (I105V/A127S/G151K/M166T/L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ.

Theo một số phương án, trình tự axit amin của biến thể Yếu tố VIII được mã hóa có độ tương đồng ít nhất là 97% với CS12-FL-AA (SEQ ID NO: 2), bao gồm năm đột biến X5 (I105V/A127S/G151K/M166T/L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ.

Theo một số phương án, trình tự axit amin của biến thể Yếu tố VIII được mã hóa có độ tương đồng ít nhất là 98% với CS12-FL-AA (SEQ ID NO: 2), bao gồm năm đột biến X5 (I105V/A127S/G151K/M166T/L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ.

Theo một số phương án, trình tự axit amin của biến thể Yếu tố VIII được mã hóa có độ tương đồng ít nhất là 99% với CS12-FL-AA (SEQ ID NO: 2), bao gồm năm đột biến X5 (I105V/A127S/G151K/M166T/L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ.

Theo một số phương án, trình tự axit amin của biến thể Yếu tố VIII được mã hóa có độ tương đồng ít nhất là 99,5% với CS12-FL-AA (SEQ ID NO: 2), bao gồm năm đột

biến X5 (I105V/A127S/G151K/M166T/L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ.

Theo một số phương án, trình tự axit amin của biến thể Yếu tố VIII được mã hóa có độ tương đồng ít nhất là 99,9% với CS12-FL-AA (SEQ ID NO: 2), bao gồm năm đột biến X5 (I105V/A127S/G151K/M166T/L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ.

Theo một số phương án, trình tự axit amin của biến thể Yếu tố VIII được mã hóa là CS12-FL-AA (SEQ ID NO: 2).

Cầu nối được thể miền B Yếu tố VIII

Như mô tả ở trên, sự liên kết giữa chuỗi nặng và chuỗi nhẹ FVIII (ví dụ như, miền B trong Yếu tố VIII kiểu đại) được làm thay đổi trong biến thể Yếu tố VIII được mô tả trong bản mô tả này. Việc loại bỏ miền B khỏi cấu trúc Yếu tố VIII kiểu đại dường như không ảnh hưởng đến hoạt tính của enzym đã hoạt hóa (ví dụ như, FVIIIa), có lẽ là vì miền B được loại bỏ trong quá trình hoạt hóa. Do ràng buộc kích thước của khả năng đóng gói AAV, các biến thể được làm khuyết miền B, được cắt cụt, và hoặc được thể cầu nối sẽ cải thiện hiệu quả của cấu trúc liệu pháp gen FVIII. Cầu nối được thể miền B thường được sử dụng nhất là cầu nối của SQ FVIII, mà chỉ giữ lại 14 axit amin của miền B làm trình tự cầu nối. Biến thể khác của VIII lợn ("OBI-1," được mô tả trong Bảng Sáng Chế Mỹ Số 6,458,563) được biểu hiện tốt ở tế bào CHO, và có cầu nối dài hơn một chút gồm 24 axit amin. Theo một số phương án, cấu trúc Yếu tố VIII được mã hóa bởi polynucleotit đã được thay đổi bộ ba mã hóa được mô tả trong bản mô tả này bao gồm trình tự cầu nối miền B kiểu SQ. Theo phương án khác, cấu trúc Yếu tố VIII được mã hóa bởi polynucleotit đã được thay đổi bộ ba mã hóa được mô tả trong bản mô tả này bao gồm trình tự cầu nối miền B kiểu OBI-1.

Mặc dù miền B Yếu tố VIII không cần thiết đối với hoạt tính, miền B chứa một vài gốc mà được cải biến sau dịch mã, ví dụ như, bằng sự glycosyl hóa liên kết N hoặc O. Phân tích in silico (Prediction of N-glycosylation sites in human proteins, R. Gupta, E. Jung và S. Brunak, in preparation (2004)) của Yếu tố VIII kiểu đại miền B dự đoán rằng ít nhất là bốn trong số các vị trí này được glycosyl hóa in vivo. Người ta cho rằng

các cải biến này ở trong miền B gây nên sự điều hòa sau dịch mã và/hoặc thời gian bán thải của Yếu tố VIII in vivo. Do đó, theo một số phương án, polypeptit cầu nối của cấu trúc Yếu tố VIII được mã hóa được mô tả trong bản mô tả này bao gồm một hoặc nhiều trình tự glycosyl hóa, để cho phép sự glycosyl hóa in vivo. Theo một số phương án, polypeptit cầu nối bao gồm ít nhất là một trình tự glycosyl hóa liên ứng (ví dụ như, trình tự liên ứng glycosyl hóa liên kết N hoặc O). Theo một số phương án, polypeptit cầu nối bao gồm ít nhất là hai trình tự glycosyl hóa liên ứng. Theo một số phương án, polypeptit cầu nối bao gồm ít nhất là ba trình tự glycosyl hóa liên ứng. Theo một số phương án, polypeptit cầu nối bao gồm ít nhất là bốn trình tự glycosyl hóa liên ứng. Theo một số phương án, polypeptit cầu nối bao gồm ít nhất là năm trình tự glycosyl hóa liên ứng. Theo một số phương án, polypeptit cầu nối bao gồm ít nhất là 6, 7, 8, 9, 10, hoặc hơn 10 trình tự glycosyl hóa liên ứng.

Theo một số phương án, polypeptit cầu nối chứa ít nhất là một trình tự glycosyl hóa được liên kết N N-X-S/T, trong đó X là axit amin bất kỳ không phải là P, S, hoặc T. Theo một số phương án, polypeptit cầu nối chứa ít nhất là hai trình tự glycosyl hóa được liên kết N N-X-S/T, trong đó X là axit amin bất kỳ không phải là P, S, hoặc T. Theo một số phương án, polypeptit cầu nối chứa ít nhất là ba trình tự glycosyl hóa được liên kết N N-X-S/T, trong đó X là axit amin bất kỳ không phải là P, S, hoặc T. Theo một số phương án, polypeptit cầu nối chứa ít nhất là bốn trình tự glycosyl hóa được liên kết N N-X-S/T, trong đó X là axit amin bất kỳ không phải là P, S, hoặc T. Theo một số phương án, polypeptit cầu nối chứa ít nhất là năm trình tự glycosyl hóa được liên kết N N-X-S/T, trong đó X là axit amin bất kỳ không phải là P, S, hoặc T. Theo một số phương án, polypeptit cầu nối chứa ít nhất là 6, 7, 8, 9, 10, hoặc hơn 10 trình tự glycosyl hóa được liên kết N N-X-S/T, trong đó X là axit amin bất kỳ không phải là P, S, hoặc T.

Theo một số phương án, polypeptit Yếu tố VIII được mã hóa được mô tả trong bản mô tả này bao gồm cầu nối miền B kiểu SQ, bao gồm các axit amin 760-762/1657-1667 của miền B Yếu tố VIII người kiểu đại (FVIII-FL-AA; SEQ ID NO:39) (Sandberg và đồng tác giả Thromb. Haemost. 85:93 (2001), nội dung của nó được kết hợp ở đây trong bản mô tả này để tham khảo). Theo một số phương án, cầu nối miền B kiểu SQ có một sự thế axit amin tương quan với trình tự kiểu đại tương ứng. Theo một số phương án, cầu nối miền B kiểu SQ có hai sự thế axit amin tương quan với trình tự kiểu đại tương ứng.

Theo một số phương án, peptit glycosyl hóa được cài xen vào trong cầu nối miền B kiểu SQ. Theo một số phương án, peptit glycosyl hóa được chọn từ các peptit glycosyl hóa được thể hiện trên Hình 8 (các SEQ ID NO: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, và 35, lần lượt, theo thứ tự xuất hiện). Theo một số phương án, peptit glycosyl hóa được mã hóa bởi polynucleotit tương ứng được thể hiện trên Hình 8 (các SEQ ID NO: 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, và 36, lần lượt, theo thứ tự xuất hiện). Theo phương án cụ thể, peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối miền B kiểu SQ của polypeptit Yếu tố VIII được mô tả trong bản mô tả này. Theo phương án cụ thể, peptit glycosyl hóa NG5 được mã hóa bởi polynucleotit có trình tự axit nucleic của SEQ ID NO:16.

Theo một số phương án, cầu nối miền B kiểu SQ chứa peptit glycosyl hóa được mã hóa bởi polynucleotit tương ứng được thể hiện trên Hình 17 (các SEQ ID NO: 40-53, lần lượt, theo thứ tự xuất hiện). Theo phương án cụ thể, cầu nối miền B kiểu SQ chứa peptit glycosyl hóa được mã hóa bởi polynucleotit có trình tự axit nucleic của SEQ ID NO:43.

Theo một số phương án, polypeptit cầu nối (ví dụ như, cầu nối miền B kiểu SQ) bao gồm peptit glycosyl hóa có độ tương đồng trình tự cao với trình tự bất kỳ trong số các SEQ ID NO: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, và 35, lần lượt, theo thứ tự xuất hiện, như được thể hiện trên các Hình 8A-8B. Theo một số phương án, peptit glycosyl hóa có không lớn hơn hai sự thể axit amin tương quan với trình tự bất kỳ trong số các SEQ ID NO: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, và 35, lần lượt, theo thứ tự xuất hiện, như được thể hiện trên các Hình 8A-8B. Theo một số phương án, peptit glycosyl hóa có không lớn hơn một sự thể axit amin tương quan với trình tự bất kỳ trong số các SEQ ID NO: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, và 35, lần lượt, theo thứ tự xuất hiện, như được thể hiện trên các Hình 8A-8B. Theo một số phương án, peptit glycosyl hóa có trình tự axit amin được chọn từ trình tự bất kỳ trong số các SEQ ID NO: 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, và 35, lần lượt, theo thứ tự xuất hiện, như được thể hiện trên các Hình 8A-8B. Theo một số phương án, peptit glycosyl hóa được mã hóa bởi trình tự polynucleotit có độ tương đồng trình tự cao với trình tự nucleotit tương ứng được chọn từ các SEQ ID NO: 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, và 36, lần lượt, theo thứ tự xuất hiện, như được thể hiện trên các Hình 8A-8B.

Theo một số phương án, peptit glycosyl hóa có không lớn hơn hai sự thể axit amin tương quan với SEQ ID NO:15 và được mã hóa bởi trình tự polynucleotit có độ tương đồng ít nhất là 90% với SEQ ID NO:16. Theo một số phương án, peptit glycosyl hóa có không lớn hơn hai sự thể axit amin tương quan với SEQ ID NO:15 và được mã hóa bởi trình tự polynucleotit có độ tương đồng ít nhất là 95% với SEQ ID NO:16. Theo một số phương án, peptit glycosyl hóa có không lớn hơn hai sự thể axit amin tương quan với SEQ ID NO:15 và được mã hóa bởi trình tự polynucleotit có độ tương đồng ít nhất là 98% với SEQ ID NO:16.

Theo một số phương án, peptit glycosyl hóa có không lớn hơn một sự thể axit amin tương quan với SEQ ID NO:15 và được mã hóa bởi trình tự polynucleotit có độ tương đồng ít nhất là 90% với SEQ ID NO:16. Theo một số phương án, peptit glycosyl hóa có không lớn hơn một sự thể axit amin tương quan với SEQ ID NO:15 và được mã hóa bởi trình tự polynucleotit có độ tương đồng ít nhất là 95% với SEQ ID NO:16. Theo một số phương án, peptit glycosyl hóa có không lớn hơn một sự thể axit amin tương quan với SEQ ID NO:15 và được mã hóa bởi trình tự polynucleotit có độ tương đồng ít nhất là 98% với SEQ ID NO:16.

Theo một số phương án, peptit glycosyl hóa có trình tự axit amin của SEQ ID NO:15 và được mã hóa bởi trình tự polynucleotit có độ tương đồng ít nhất là 90% với SEQ ID NO:16. Theo một số phương án, peptit glycosyl hóa có trình tự axit amin của SEQ ID NO:15 và được mã hóa bởi trình tự polynucleotit có độ tương đồng ít nhất là 95% với SEQ ID NO:16. Theo một số phương án, peptit glycosyl hóa có trình tự axit amin của SEQ ID NO:15 và được mã hóa bởi trình tự polynucleotit có độ tương đồng ít nhất là 98% với SEQ ID NO:16.

Các thành phần điều hòa *cis*

Theo một số phương án, hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII, như nêu trong bản mô tả này, còn bao gồm một hoặc nhiều thành phần điều hòa tác động *cis*, ví dụ như vùng khởi động và/hoặc thành phần vùng tăng cường, mà gây ra sự biểu hiện gen *in vivo*, mà được liên kết có điều khiển với polynucleotit mã hóa cho biến thể Yếu tố VIII. Trong ngữ cảnh này, “tác động *cis*” có nghĩa là thành phần điều hòa có mặt trên cùng phân tử ADN như gen mà chúng điều hòa. Như người có trình độ trung bình trong lĩnh vực hiểu rõ và được thảo luận dưới đây, các thành phần điều hòa thích hợp để

sử dụng trong sáng chế bao gồm, nhưng không giới hạn ở, vùng khởi động, thành phần vùng tăng cường, thành phần tín hiệu polyadenyl hóa, và/hoặc thành phần đoạn lặp đầu cuối đảo ngược.

Vùng khởi động

Vùng khởi động để sử dụng trong sáng chế được liên kết có điều khiển với vùng mã hóa, thường là được liên kết trực tiếp (ví dụ như không có nucleotit nào nữa được bao gồm giữa vùng khởi động và vùng mã hóa) mặc dù theo một số phương án, sự liên kết gián tiếp có thể được sử dụng, ví dụ như thông qua việc sử dụng cầu nối không có chức năng, hoặc trong trường hợp mà các thành phần điều hòa bổ sung mà hoạt động “xuôi dòng” của vùng khởi động nhưng “ngược dòng” của vùng mã hóa có thể được sử dụng. Tuy nhiên, do sự ràng buộc kích thích của vectơ theo sáng chế, nó thường không được ưu tiên.

Thành phần vùng tăng cường

Theo một số phương án, một hoặc nhiều thành phần vùng tăng cường được sử dụng trong cấu trúc biến thể Yếu tố VIII. Như đã biết trong lĩnh vực, thành phần vùng tăng cường thường gây ra sự biểu hiện theo cách phụ thuộc mô, ví dụ như, chủ yếu trong mô đặc hiệu. Nhìn chung, như được mô tả dưới đây, thành phần vùng tăng cường thường nằm “ngược dòng” của thành phần vùng khởi động. Vì Yếu tố VIII được tổng hợp chủ yếu trong gan, theo một số phương án, hợp phần axit nucleic được mô tả trong bản mô tả này bao gồm thành phần điều hòa đặc hiệu gan, mà về cơ bản giới hạn sự biểu hiện của biến thể Yếu tố VIII được mã hóa ở tế bào gan.

Thông thường, các thành phần điều hòa đặc hiệu gan có thể bắt nguồn từ gen đã biết bất kỳ để được biểu hiện riêng ở trong gan. WO 2009/130208 xác định một vài gen được biểu hiện theo kiểu đặc hiệu gan, bao gồm, chất ức chế serpin peptidaza, thành viên nhánh A 1, còn gọi là α -antitrypsin (SERPINA1; GeneID 5265), apolipoprotein C-I (APOC1; GeneID 341), apolipoprotein C-IV (APOC4; GeneID 346), apolipoprotein H (APOH; GeneID 350); transthyretin (TTR; GeneID 7276), albumin (ALB; GeneID 213), aldolaza B (ALDOB; GeneID 229), xytochrom P450, họ 2, phân họ E, polypeptit 1 (CYP2E1; GeneID 1571), chuỗi alpha fibrinogen (FGA; GeneID 2243), transferin (TF; GeneID 7018), protein liên quan đến haptoglobin (HPR; GeneID 3250). Theo một số phương án, hợp phần axit nucleic được mô tả trong bản mô tả này bao gồm thành

phần điều hòa đặc hiệu gan bắt nguồn từ locut hệ gen của một hoặc nhiều protein này. Một vài ví dụ về các thành phần này được mô tả trong WO 2009/130208, nội dung của nó được kết hợp rõ ràng trong bản mô tả này để tham khảo, đến toàn bộ nội dung của nó, cho mọi mục đích.

Một ví dụ về thành phần điều hòa đặc hiệu gan là từ gen transthyretin (TTR), được đề cập chung dưới dạng "TTR_e" hoặc "TTR_{Enh}." Hsieh J.L., và đồng tác giả, Cancer Sci., 100(3):537-45 (2009), nội dung của nó được kết hợp rõ ràng trong bản mô tả này để tham khảo, đến toàn bộ nội dung của nó, cho mọi mục đích. Theo một số phương án, hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII, như được mô tả trong bản mô tả này, bao gồm thành phần vùng khởi động TTR người. Theo một phương án, vùng khởi động TTR người có trình tự axit nucleic có độ tương đồng trình tự cao với vùng khởi động hTTR được thể hiện trên Hình 4 (SEQ ID NO:6). Theo một số phương án, vùng khởi động TTR người có trình tự axit nucleic mà tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100% với SEQ ID NO:6.

Ví dụ khác về thành phần điều hòa đặc hiệu gan là từ gen SERPINA1, như được mô tả trong Công Bố PCT Số WO 2016/146757, nội dung của nó được kết hợp rõ ràng trong bản mô tả này để tham khảo, đến toàn bộ nội dung của nó, cho mọi mục đích. Ví dụ về thành phần này là thành phần điều khiển điều hòa CRM8 được thể hiện trên Hình 4 (SEQ ID NO:5). Theo một số phương án, thành phần điều khiển điều hòa điều khiển SERPINA1 có trình tự axit nucleic mà tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, hoặc 100% với CRM8 (SEQ ID NO:5).

Theo một số phương án, hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII, như nêu trong bản mô tả này, bao gồm một hoặc nhiều thành phần điều khiển điều hòa điều khiển SERPINA1, như được lấy ví dụ bằng các cấu trúc được minh họa trên Hình 11. Theo một phương án, polynucleotit Yếu tố VIII bao gồm một thành phần điều khiển điều hòa điều khiển SERPINA1 (ví dụ như, CRM8). Theo phương án khác, polynucleotit Yếu tố VIII bao gồm hai thành phần điều khiển điều hòa điều khiển SERPINA1 (ví dụ như, CRM8). Theo phương án khác nữa, polynucleotit Yếu tố VIII bao gồm 3, 4, 5, 6, hoặc hơn 6 thành phần điều khiển điều hòa điều khiển SERPINA1 (ví dụ như, CRM8).

Theo một phương án, hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII, như nêu trong bản mô tả này, bao gồm một hoặc nhiều thành phần điều khiển điều hòa điều khiển SERPINA1 (ví dụ như, CRM8) và thành phần vùng khởi động TTR người, như được lấy ví dụ trên Hình 11. Theo một phương án, hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII, như nêu trong bản mô tả này, bao gồm hai thành phần CRM8 và thành phần vùng khởi động TTR người được liên kết có điều khiển với polynucleotit mã hóa cho biến thể Yếu tố VIII.

Theo một số phương án, hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII, như nêu trong bản mô tả này, bao gồm một hoặc nhiều thành phần CRM8 nằm ngược dòng của vùng khởi động TTR người. Ví dụ như, một hoặc nhiều thành phần CRM8 nằm ở 5' của vùng khởi động TTR trong cấu trúc sợi kép, tương quan với chiều phiên mã của phân tử. Điều này có nghĩa là trong cấu trúc sợi đơn (+) (ví dụ như, trong đó sợi đơn mã hóa cho biến thể Yếu tố VIII), một hoặc nhiều thành phần CRM8 nằm ở 5' của vùng khởi động TTR.

Như báo cáo trong Ví dụ 6, do kích thước lớn của cấu trúc biến thể Yếu tố VIII được mô tả trong bản mô tả này, sự giảm đi nhỏ của số lượng nucleotit trong polynucleotit Yếu tố VIII mà là một phần của vector liệu pháp gen AAV có thể làm tăng đáng kể hiệu lực sinh học Yếu tố VIII của vector. Theo đó, theo một số phương án, một hoặc nhiều thành phần CRM8 được gắn trực tiếp vào đầu 5' của vùng khởi động TTR, ví dụ như, không có các nucleotit ngoại lai nằm giữa thành phần CRM8 và vùng khởi động TTR. Tương tự, theo một số phương án, vùng khởi động TTR được gắn trực tiếp vào đầu 5' của trình tự mã hóa, hoặc vào trình tự mở đầu dịch mã (ví dụ như, trình tự Kozak), đối với polypeptit biến thể Yếu tố VIII, ví dụ như, không có các nucleotit ngoại lai nằm giữa vùng khởi động TTR và gen biến thể Yếu tố VIII.

Tín hiệu polyadenyl hóa

Theo một số phương án, thành phần điều hòa đối với các cấu trúc được mô tả trong bản mô tả này (ví dụ như hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII) là thành phần điều hòa mà là tín hiệu polyadenyl hóa, ví dụ như, như được minh họa trong các cấu trúc ví dụ trên Hình 11. Tín hiệu polyadenyl hóa định hướng sự tổng hợp của đuôi poly-A trên đầu 3' của bản phiên mã mRNA được tạo ra từ polynucleotit Yếu tố VIII. Theo đó, tín hiệu polyadenyl hóa nằm ở 3' đối với trình tự mã hóa biến thể Yếu tố VIII.

tổ VIII. Ví dụ không làm giới hạn sáng chế về tín hiệu polyadenyl hóa mà có thể được sử dụng trong cấu trúc liệu pháp gen Yếu tố IX được mô tả trong bản mô tả này bao gồm tín hiệu polyadenyl hóa tổng hợp, tín hiệu poly-adenyl hóa bắt nguồn từ gen muộn virus khỉ 40 (SV40), tín hiệu polyadenyl hóa hormon sinh trưởng bò (BGH), và tín hiệu polyadenyl hóa gen β -globin thỏ nhỏ nhất (mRBG).

Theo một số phương án, hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII, như nêu trong bản mô tả này, bao gồm tín hiệu polyadenyl hóa tổng hợp, như được lấy ví dụ bằng các cấu trúc được minh họa trên Hình 11. Theo một phương án, tín hiệu polyadenyl hóa tổng hợp có trình tự axit nucleic mà tương đồng ít nhất là 90%, 95%, 97%, hoặc 100% với tín hiệu Poly-A tổng hợp được thể hiện trên Hình 4 (SEQ ID NO:8).

Như báo cáo trong Ví dụ 6, do kích thước lớn của cấu trúc biến thể Yếu tố VIII được mô tả trong bản mô tả này, sự giảm đi nhỏ của số lượng nucleotit trong polynucleotit Yếu tố VIII mà là một phần của vector liệu pháp gen AAV có thể làm tăng đáng kể hiệu lực sinh học Yếu tố VIII của vector. Theo đó, theo một số phương án, tín hiệu polyadenyl hóa được gắn trực tiếp vào đầu 3' của trình tự mã hóa của polypeptit biến thể Yếu tố VIII, bao gồm một hoặc nhiều bộ ba kết thúc mà nằm ở đầu của trình tự mã hóa. Ví dụ như, không có các nucleotit ngoại lai nằm giữa biến thể Yếu tố VIII gen và trình tự polyadenyl hóa.

Đoạn lặp đầu cuối đảo ngược

Theo một số phương án, hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII, như nêu trong bản mô tả này, còn bao gồm các đoạn lặp đầu cuối đảo ngược (ITR) virus kết hợp adeno (AAV) kẹp hai bên trình tự mã hóa biến thể Yếu tố VIII và các thành phần điều hòa kết hợp (ví dụ như, vùng khởi động, vùng tăng cường, tín hiệu polyadenyl hóa, v.v.). Mỗi đoạn lặp đầu cuối đảo ngược tạo thành kẹp tóc, mà làm thuận lợi cho việc tự môi cho sự tổng hợp không phụ thuộc primaza của sợi ADN thứ hai. Các ITR cũng giúp làm thuận lợi cho sự bao nang ở trong virion AAV và sự tích hợp của hệ gen AAV vào trong hệ gen tế bào chủ.

Theo một số phương án, hợp phần axit nucleic mã hóa cho biến thể Yếu tố VIII, như nêu trong bản mô tả này, bao gồm 5' ITR có độ tương đồng trình tự cao (ví dụ như, tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, hoặc 100%) với AAV2 5' ITR được thể hiện trên Hình 4 (SEQ ID NO:4). Theo một số phương án, hợp phần axit

nucleic mã hóa cho biến thể Yếu tố VIII, như nêu trong bản mô tả này, bao gồm 3' ITR có độ tương đồng trình tự cao (ví dụ như, tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, hoặc 100%) với AAV2 3' ITR được thể hiện trên Hình 4 (SEQ ID NO:9).

Như báo cáo trong Ví dụ 6, do kích thước lớn của cấu trúc biến thể Yếu tố VIII được mô tả trong bản mô tả này, sự giảm đi nhỏ của số lượng nucleotit trong polynucleotit Yếu tố VIII mà là một phần của vector liệu pháp gen AAV có thể làm tăng đáng kể hiệu lực sinh học Yếu tố VIII của vector. Theo đó, theo một số phương án, 5' ITR được gắn trực tiếp vào đầu 5' của thành phần đặc hiệu gan (ví dụ như, một hoặc nhiều thành phần CRM8), sao cho không có nucleotit ngoại lai nào nằm giữa trình tự 5' ITR và thành phần đặc hiệu gan. Tương tự, theo một số phương án, 3' ITR được gắn trực tiếp vào đầu 3' của tín hiệu polyadenyl hóa, sao cho không có nucleotit ngoại lai nào nằm giữa tín hiệu polyadenyl hóa và trình tự 3' ITR.

IV. Vector biểu hiện Yếu tố VIII

Theo một số phương án, polynucleotit đã được thay đổi bộ ba mã hóa được mô tả trong bản mô tả này được tích hợp vào vector biểu hiện. Ví dụ không làm giới hạn sáng chế về vector biểu hiện bao gồm vector virus (ví dụ như, vector thích hợp cho liệu pháp gen), vector plasmid, vector thể thực khuẩn, cosmid, phagemid, nhiễm sắc thể nhân tạo, và dạng tương tự. Nhìn chung, có hai loại cơ bản của vector biểu hiện để sử dụng trong sáng chế: các vector biểu hiện mà được sử dụng trong môi trường nuôi cấy tế bào để sản xuất polypeptit Yếu tố VIII, và vector biểu hiện mà được sử dụng làm vector liệu pháp gen để dùng cho bệnh nhân sao cho hàm lượng nội sinh của Yếu tố VIII (dù là protein hay hoạt tính) được tăng lên.

Ví dụ không làm giới hạn sáng chế về vector virus bao gồm: retrovirus, ví dụ như, virus gây bệnh bạch cầu ở chuột Moloney (MMLV), virus gây sarcoma ở chuột Harvey, virus khối u vú ở chuột, và virus gây sarcoma Rous; adenovirus, virus kết hợp adeno; virus kiểu SV40; polyomavirus; virus Epstein-Barr; virus u nhú; virus herpes; virus đậu mùa; và virus bại liệt.

Theo nhiều phương án, polynucleotit được tối ưu hóa bộ ba mã hóa theo sáng chế được sử dụng trong các ứng dụng liệu pháp gen, sao cho việc dùng cho bệnh nhân dẫn đến sự sản xuất của Yếu tố VIII như được mô tả chung trong bản mô tả này. Nhìn chung,

vector virus liệu pháp gen tốt hơn là không có khả năng sao chép, sao cho sự đưa vào của vector liệu pháp gen vào bệnh nhân không dẫn đến sự nhân lên của virus.

Theo đó, theo một số phương án, polynucleotit đã được thay đổi bộ ba mã hóa được mô tả trong bản mô tả này được tích hợp vào vector liệu pháp gen. Theo một số phương án, vector liệu pháp gen là retrovirus, và cụ thể là retrovirus không có khả năng sao chép. Theo một số phương án, polynucleotit đã được thay đổi bộ ba mã hóa được mô tả trong bản mô tả này được tích hợp vào vector biểu hiện retrovirus. Các hệ thống này đã được mô tả trước đây, và thường được biết rõ trong lĩnh vực (Mann và đồng tác giả, *Cell*, 33:153-159, 1983; Nicolas và Rubinstein, In: *Vectors: A survey of molecular cloning vectors and their uses*, Rodriguez và Denhardt, eds., Stoneham: Butterworth, pp. 494-513, 1988; Temin, In: *Gene Transfer*, Kucherlapati (ed.), New York: Plenum Press, pp. 149-188, 1986). Theo phương án cụ thể, vector retrovirus là vector lentivirus (xem, ví dụ như, Naldini và đồng tác giả, *Science*, 272(5259):263-267, 1996; Zufferey và đồng tác giả, *Nat Biotechnol*, 15(9):871-875, 1997; Blomer và đồng tác giả, *J Virol.*, 71(9):6641-6649, 1997; Bằng Sáng Chế Mỹ Số 6,013,516 và 5,994,136). Quy trình sản xuất retrovirus không có khả năng sao chép đã được biết đến trong lĩnh vực. Về tổng quan, xem Kriegler, M., *Gene Transfer and Expression, A Laboratory Manual*, W.H. Freeman Co., New York (1990) và Murry, E. J., *Methods in Molecular Biology*, Vol. 7, Humana Press, Inc., Clifton, N.J. (1991).

Nhiều loại vector có thể được sử dụng để biểu hiện polypeptit Yếu tố VIII từ polypeptit đã được thay đổi bộ ba mã hóa trong môi trường nuôi cấy tế bào, bao gồm vector biểu hiện sinh vật nhân chuẩn và sinh vật nhân sơ. Theo các phương án nhất định, vector plasmit được dự tính để sử dụng trong việc biểu hiện polypeptit Yếu tố VIII trong môi trường nuôi cấy tế bào. Nhìn chung, vector plasmit chứa đơn vị sao chép và các trình tự điều khiển mà bắt nguồn từ các loài tương thích với tế bào chủ được sử dụng liên quan đến các vật chủ này. Vector có thể mang vị trí sao chép, cũng như là các trình tự đánh dấu mà có khả năng mang lại sự chọn lọc kiểu hình trong tế bào được biến nạp. Plasmit bao gồm polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho polypeptit Yếu tố VIII, được liên kết có điều khiển với một hoặc nhiều trình tự điều khiển, ví dụ như, vùng khởi động.

Ví dụ không làm giới hạn sáng chế về vector cho sự biểu hiện ở sinh vật nhân sơ bao gồm plasmit chẳng hạn như pRSET, pET, pBAD, v.v., trong đó vùng khởi động dùng trong vector biểu hiện sinh vật nhân sơ bao gồm lac, trc, trp, recA, araBAD, v.v.. Ví dụ về vector cho sự biểu hiện ở sinh vật nhân chuẩn bao gồm: (i) để biểu hiện ở nấm men, vector chẳng hạn như pAO, pPIC, pYES, pMET, bằng cách sử dụng vùng khởi động chẳng hạn như AOX1, GAP, GAL1, AUG1, etc; (ii) để biểu hiện ở tế bào côn trùng, vector chẳng hạn như pMT, pAc5, pIB, pMIB, pBAC, v.v., bằng cách sử dụng vùng khởi động chẳng hạn như PH, p10, MT, Ac5, OpIE2, gp64, polh, v.v., và (iii) để biểu hiện trong tế bào động vật có vú, vector chẳng hạn như pSVL, pCMV, pRc/RSV, pcDNA3, pBPV, v.v., và vector bắt nguồn từ hệ thống virus chẳng hạn như virus đậu mùa, virus kết hợp adeno, virus herpes, retrovirus, v.v., bằng cách sử dụng vùng khởi động chẳng hạn như CMV, SV40, EF-1, UbC, RSV, ADV, BPV, và β -actin.

Vector virus kết hợp adeno (AAV)

Theo một phương án, polynucleotit đã được thay đổi bộ ba mã hóa, như được mô tả trong bản mô tả này, được tích hợp vào vector liệu pháp gen dựa trên virus kết hợp adeno (AAV). Hệ thống AAV đã được mô tả trước đây và thường đã được biết rõ trong lĩnh vực (Kelleher và Vos, *Biotechniques*, 17(6):1110-17 (1994); Cotten và đồng tác giả, *Proc Natl Acad Sci USA*, 89(13):6094-98 (1992); Curiel, *Nat Immun*, 13(2-3):141-64 (1994); Muzyczka, *Curr Top Microbiol Immunol*, 158:97-129 (1992); và Asokan A, và đồng tác giả, *Mol. Ther.*, 20(4):699-708 (2012), mỗi tài liệu này được kết hợp trong bản mô tả này để tham khảo trong toàn bộ nội dung của chúng cho mọi mục đích). Chi tiết về việc tạo ra và việc sử dụng rvector AAV được mô tả, ví dụ như, trong các Bằng Sáng Chế Mỹ Số 5,139,941 và 4,797,368, mỗi tài liệu này được kết hợp trong bản mô tả này để tham khảo trong toàn bộ nội dung của chúng cho mọi mục đích.

Theo đó, theo một số phương án, vector liệu pháp gen AAV được đề xuất mà bao gồm polynucleotit đã được thay đổi bộ ba mã hóa (ví dụ như, hợp phần axit nucleic), như được mô tả trong bản mô tả này, mà bao gồm trình tự nucleotit có độ tương đồng trình tự cao với các phần của CS12-FL-NA (SEQ ID NO:1) mã hóa cho chuỗi nặng và chuỗi nhẹ Yếu tố VIII, và cầu nối được thể miền B có 14 axit amin, ngắn (ví dụ như, cầu nối "SQ" chứa vị trí phân cắt furin để làm thuận lợi cho sự trưởng thành của protein FVIIIa hoạt tính in vivo), và còn bao gồm một hoặc nhiều trong số năm đột biến X5 (ví

dụ như, một, hai, ba, bốn, hoặc tất cả năm của đột biến I105V, A127S, G151K, M166T, L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và/hoặc peptit glycosyl hóa ngắn (ví dụ như, NG5; SEQ ID NO:15) được cài xen vào cầu nối được thể miền B (ví dụ như, cầu nối SQ).

Theo một số phương án, vector liệu pháp gen AAV bao gồm polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa mà có trình tự nucleotit có độ tương đồng trình tự cao với CS12-FL-NA (SEQ ID NO:1) và mã hóa cho polypeptit Yếu tố VIII có chuỗi nặng và chuỗi nhẹ Yếu tố VIII người, và cầu nối được thể miền B có 14 axit amin, ngắn (cầu nối “SQ”) chứa vị trí phân cắt furin để làm thuận lợi cho sự trưởng thành của protein FVIIIa hoạt tính in vivo, trong đó chuỗi nặng của polypeptit Yếu tố VIII bao gồm năm đột biến X5 (I105V, A127S, G151K, M166T, và L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ. Theo một số phương án, polynucleotit Yếu tố VIII có trình tự nucleotit có độ tương đồng ít nhất là 95%, 96%, 97%, 98%, 99%, 99,5%, 99,9%, hoặc 100% với CS12-FL-NA (SEQ ID NO:1).

Theo một số phương án, biến thể Yếu tố VIII được mã hóa bởi polynucleotit đã được thay đổi bộ ba mã hóa của vector liệu pháp gen AAV có trình tự axit amin có độ tương đồng trình tự cao với CS12-FL-AA (SEQ ID NO: 2), bao gồm năm đột biến X5 (I105V/A127S/G151K/M166T/L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ. Theo một số phương án, trình tự axit amin của biến thể Yếu tố VIII được mã hóa có độ tương đồng ít nhất là 97%, 98%, 99%, 99,5%, 99,9%, hoặc 100% với CS12-FL-AA (SEQ ID NO: 2), bao gồm năm đột biến X5 (I105V/A127S/G151K/M166T/L171P (SPI), tương quan với trình tự Yếu tố VIII kiểu đại người chiều dài đầy đủ), và peptit glycosyl hóa NG5 (SEQ ID NO:15) được cài xen vào trong cầu nối SQ.

Theo một số phương án, polynucleotit mã hóa cho biến thể Yếu tố VIII ở trong vector liệu pháp gen AAV được liên kết có điều khiển với thành phần vùng khởi động có trình tự axit nucleic có độ tương đồng trình tự cao với vùng khởi động hTTR được thể hiện trên Hình 4 (SEQ ID NO:6). Theo một số phương án, thành phần vùng khởi

động có trình tự axit nucleic mà tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100% với SEQ ID NO:6.

Theo một số phương án, polynucleotit mã hóa cho biến thể Yếu tố VIII ở trong vector liệu pháp gen AAV được liên kết có điều khiển với một hoặc nhiều thành phần điều hòa đặc hiệu gan có trình tự axit nucleic có độ tương đồng trình tự cao với thành phần CRM8 được thể hiện trên Hình 4 (SEQ ID NO:5). Theo một số phương án, các thành phần điều hòa đặc hiệu gan có trình tự axit nucleic mà tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, 99,5%, hoặc 100% với SEQ ID NO:5. Theo một số phương án, như được minh họa trên Hình 11, polynucleotit bao gồm một thành phần CRM8. Theo một số phương án, như được minh họa trên Hình 11, polynucleotit bao gồm hai thành phần CRM.

Theo một số phương án, polynucleotit mã hóa cho biến thể Yếu tố VIII ở trong vector liệu pháp gen AAV được liên kết có điều khiển với một thành phần CRM8 và thành phần vùng khởi động TTR người, như được lấy ví dụ trên Hình 11. Theo một số phương án, polynucleotit mã hóa cho biến thể Yếu tố VIII ở trong vector liệu pháp gen AAV được liên kết có điều khiển với hai thành phần CRM8 và thành phần vùng khởi động TTR người, như được lấy ví dụ trên Hình 11. Như được mô tả trong Ví dụ 2, việc sử dụng vùng khởi động hTTR và một hoặc hai thành phần CRM8 đặc hiệu gan làm tăng hiệu lực sinh học Yếu tố VIII ngoại sinh in vivo trong tế bào HepG2 lần lượt lên khoảng 2 lần và 4 lần, khi so với việc sử dụng các trình tự vùng khởi động và vùng tăng cường TTR chuột nhất (so sánh vCS115 và vCS116 với vCS04 trên Hình 12).

Như báo cáo trong Ví dụ 6, do kích thước lớn của cấu trúc biến thể Yếu tố VIII được mô tả trong bản mô tả này, sự giảm đi nhỏ của số lượng nucleotit trong polynucleotit Yếu tố VIII mà là một phần của vector liệu pháp gen AAV có thể làm tăng đáng kể hiệu lực sinh học Yếu tố VIII của vector. Theo đó, theo một số phương án, một hoặc nhiều thành phần CRM8 được gắn trực tiếp vào đầu 5' của vùng khởi động TTR, ví dụ như, không có các nucleotit ngoại lai nằm giữa thành phần CRM8 và vùng khởi động TTR. Tương tự, theo một số phương án, vùng khởi động TTR được gắn trực tiếp vào đầu 5' của trình tự mã hóa, hoặc vào trình tự mở đầu dịch mã (ví dụ như, trình tự Kozak), đối với polypeptit biến thể Yếu tố VIII, ví dụ như, không có các nucleotit ngoại lai nằm giữa vùng khởi động TTR và gen biến thể Yếu tố VIII.

Theo một số phương án, polynucleotit mã hóa cho biến thể Yếu tố VIII ở trong vector liệu pháp gen AAV được liên kết có điều khiển với tín hiệu polyadenyl hóa, ví dụ như, như được minh họa trong các cấu trúc ví dụ trên Hình 11. Theo một phương án, tín hiệu polyadenyl hóa tổng hợp có trình tự axit nucleic mà tương đồng ít nhất là 90%, 95%, 97%, hoặc 100% với tín hiệu Poly-A tổng hợp được thể hiện trên Hình 4 (SEQ ID NO:8).

Như báo cáo trong Ví dụ 6, do kích thước lớn của cấu trúc biến thể Yếu tố VIII được mô tả trong bản mô tả này, sự giảm đi nhỏ của số lượng nucleotit trong polynucleotit Yếu tố VIII mà là một phần của vector liệu pháp gen AAV có thể làm tăng đáng kể hiệu lực sinh học Yếu tố VIII của vector. Theo đó, theo một số phương án, tín hiệu polyadenyl hóa được gắn trực tiếp vào đầu 3' của trình tự mã hóa của polypeptit biến thể Yếu tố VIII, bao gồm một hoặc nhiều bộ ba kết thúc mà nằm ở đầu của trình tự mã hóa. Ví dụ như, không có các nucleotit ngoại lai nằm giữa biến thể Yếu tố VIII gen và trình tự polyadenyl hóa.

Đoạn lặp đầu cuối đảo ngược là các thành phần điều hòa cis cần thiết cho vector tái tổ hợp dựa trên AAV. Theo đó, polynucleotit mã hóa cho biến thể Yếu tố VIII được sử dụng trong vector liệu pháp gen AAV được mô tả trong bản mô tả này bao gồm các trình tự 5' và 3' ITR. Theo một số phương án, 5' ITR có độ tương đồng trình tự cao (ví dụ như, tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, hoặc 100%) với AAV2 5' ITR được thể hiện trên Hình 4 (SEQ ID NO:4). Theo một số phương án, 3' ITR có độ tương đồng trình tự cao (ví dụ như, tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, hoặc 100%) với AAV2 3' ITR được thể hiện trên Hình 4 (SEQ ID NO:9).

Như báo cáo trong Ví dụ 6, do kích thước lớn của cấu trúc biến thể Yếu tố VIII được mô tả trong bản mô tả này, sự giảm đi nhỏ của số lượng nucleotit trong polynucleotit Yếu tố VIII mà là một phần của vector liệu pháp gen AAV có thể làm tăng đáng kể hiệu lực sinh học Yếu tố VIII của vector. Theo đó, theo một số phương án, 5' ITR được gắn trực tiếp vào đầu 5' của thành phần đặc hiệu gan (ví dụ như, một hoặc nhiều thành phần CRM8), sao cho không có nucleotit ngoại lai nào nằm giữa trình tự 5' ITR và thành phần đặc hiệu gan. Tương tự, theo một số phương án, 3' ITR được gắn trực tiếp vào đầu 3' của tín hiệu polyadenyl hóa, sao cho không có nucleotit ngoại lai nào nằm giữa tín hiệu polyadenyl hóa và trình tự 3' ITR.

Theo phương án cụ thể, như được lấy ví dụ trên Hình 11, polynucleotit được bao gồm ở trong vector liệu pháp gen AAV được đề xuất trong bản mô tả này bao gồm, theo chiều từ 5' đến 3', trình tự 5' ITR (ví dụ như, có độ tương đồng trình tự cao với SEQ ID NO:4), một hoặc hai thành phần CRM có độ tương đồng trình tự cao với SEQ ID NO:5, thành phần vùng khởi động hTTR có độ tương đồng trình tự cao với SEQ ID NO:6, trình tự liên ứng Kozak nhỏ nhất có độ tương đồng trình tự cao với SEQ ID NO:7, polynucleotit biến thể Yếu tố VIII có độ tương đồng trình tự cao với SEQ ID NO:1, trình tự polyadenyl hóa (ví dụ như, có độ tương đồng trình tự cao với SEQ ID NO: 8), và trình tự 3' AAV ITR (ví dụ như, có độ tương đồng trình tự cao với SEQ ID NO:9). Theo một số phương án, polynucleotit được mô tả có độ tương đồng trình tự cao (ví dụ như, tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, 99,5%, 99,9%, hoặc 100%) với vector CS12-CRM8.2-Vr được thể hiện trên Hình 3 (SEQ ID NO:3).

Vector liệu pháp gen AAV được mô tả trong bản mô tả này được sử dụng với protein capsit AAV mà bao nang polynucleotit mã hóa cho polypeptit biến thể Yếu tố VIII, như được mô tả trong bản mô tả này. Tức là, sự phân phối của vector virus mà sẽ sản xuất ra Yếu tố VIII được thực hiện bằng cách sử dụng hạt virus bao gồm “vỏ” protein capsit mà bao nang vector virus.

Typ huyết thanh của vector AAV thường được xác định bằng protein capsit được sử dụng. Một vài typ huyết thanh AAV đã được xác định đặc điểm, bao gồm AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, và AAV9. Thông thường, typ huyết thanh AAV bất kỳ có thể được sử dụng cho cấu trúc liệu pháp gen Yếu tố VIII được mô tả trong bản mô tả này. Tuy nhiên, các typ huyết thanh có tính hướng khác nhau, ví dụ như, chúng ưu tiên nhiễm các mô khác nhau. Theo một phương án, bởi vì Yếu tố VIII được sản xuất chủ yếu trong gan, typ huyết thanh AAV đối với cấu trúc liệu pháp gen được bộc lộ được chọn dựa trên tính hướng gan, được phát hiện thấy trong ít nhất là các typ huyết thanh AAV7, AAV8, và AAV9. Theo đó, trong một phương án, cấu trúc liệu pháp gen Yếu tố VIII là vector typ huyết thanh AAV7. Theo phương án khác, cấu trúc liệu pháp gen Yếu tố VIII là vector typ huyết thanh AAV8. Theo phương án khác nữa, cấu trúc liệu pháp gen Yếu tố VIII là vector typ huyết thanh AAV9.

Theo một số phương án, polynucleotit plasmit mà kết hợp hệ gen liệu pháp gen Yếu tố VIII đã được thay đổi bộ ba mã hóa cũng được đề xuất. Các plasmit này hữu

dụng để sản xuất hạt AAV cuối cùng (ví dụ như, virion AAV mang polynucleotit mã hóa cho polypeptit Yếu tố VIII biến thể), ví dụ như, khi được đưa vào tế bào động vật có vú có thẩm quyền đối với sự sản xuất AAV tái tổ hợp (ví dụ như, tế bào chứa axit nucleic mã hóa cho gen *rep* và *cap* AAV, cũng như là gen giúp đỡ (ví dụ như, gen adenovirut) để sản xuất AAV. Theo một số phương án, plasmit bao gồm các thành phần điều hòa mà cho phép sự sao chép của plasmit (ví dụ như, để tăng quy mô plasmit) trong tế bào chủ (ví dụ như, tế bào chủ nhân sơ, chẳng hạn như vi khuẩn, hoặc tế bào chủ nhân chuẩn, chẳng hạn như nấm men).

Ví dụ như, trình tự của plasmit ví dụ mang hệ gen liệu pháp gen Yếu tố VIII đã được thay đổi bộ ba mã hóa được thể hiện trên Hình 5 (CS12-CRM8.2-Vrp; SEQ ID NO:10), theo một phương án của sáng chế. Plasmit CS12-CRM8.2-Vrp bao gồm hệ gen liệu pháp gen Yếu tố VIII CS12-CRM8.2-Vr (được thể hiện dưới dạng SEQ ID NO:3, trên Hình 3), và khung plasmit (được thể hiện dưới dạng SEQ ID NO:54, trên Hình 18). Các thành phần di truyền của hệ gen liệu pháp gen Yếu tố VIII CS12-CRM8.2-Vr được thể hiện trên Hình 4, như được mô tả ở trên. Khung plasmit của plasmit CS12-CRM8.2-Vrp bao gồm đơn vị sao chép pMB1 (được thể hiện dưới dạng SEQ ID NO:55 trên Hình 19; Bolívar F., Life Sci., 25(10):807-17 (1979)) mà làm thuận lợi cho sự sao chép của plasmit trong tế bào chủ vi khuẩn, và gen kháng ampicillin Bla(ApR) (được thể hiện dưới dạng SEQ ID NO:56 trên Hình 19; Sutcliffe, P.N.A.S. U.S.A, 75(8):3737-41 (1978)) mà làm thuận lợi cho sự chọn lọc tế bào chủ vi khuẩn được biến nạp bằng plasmit. Vị trí của mỗi thành phần trong plasmit CS12-CRM8.2-Vrp được thể hiện trong Bảng 1, dưới đây.

Bảng 1. Các thành phần có mặt trong plasmit CS12-CRM8.2-Vrp.

Tên của thành phần	Vị trí nucleotit	SEQ ID NO:
AAV2 5'-ITR	1 - 145	SEQ ID NO:4
CRM8	146 – 217; 219 – 290	SEQ ID NO:5
Vùng khởi động TTR người	291 - 523	SEQ ID NO:6
Trình tự Kozak	524 - 528	SEQ ID NO:7
Trình tự mã hóa FVIII-BDD với X5 và NG5	529 - 4923	SEQ ID NO:1
Bộ ba mở đầu ATG	529 - 531	
Bộ ba kết thúc TGA	4921 – 4923; 4924 - 4926	
Biến thể X5 (I86V; A108S; G132K; M147T; L152P)	841 – 843; 907 – 909; 979 – 981; 1024 – 1026; 1039 - 1041	

Trình tự NG5	2821 - 2841	SEQ ID NO:16
PolyA tổng hợp	4927 - 4975	SEQ ID NO:8
AAV2 3'-ITR	4976 - 5120	SEQ ID NO:9
Khung plasmit	5121 – 7794	SEQ ID NO:54
☐ hRep (pMB1)	☐ 15541 – 6155	SEQ ID NO:55
☐ eBla(ApR)	☐ 66315 - 7175	SEQ ID NO:56

Theo một phương án, plasmit kết hợp hệ gen liệu pháp gen Yếu tố VIII biến thể có độ tương đồng trình tự cao với plasmit CS12-CRM8.2-Vrp được thể hiện trên Hình 5 (SEQ ID NO:10). Theo một số phương án, polynucleotit được mô tả có độ tương đồng trình tự cao (ví dụ như, tương đồng ít nhất là 90%, 95%, 96%, 97%, 98%, 99%, 99,5%, 99,9%, hoặc 100%) với plasmit CS12-CRM8.2-Vrp được thể hiện trên Hình 5 (SEQ ID NO:10). Theo một số phương án, polypeptit được mô tả bao gồm trình tự mã hóa Yếu tố VIII-BDD, ví dụ như, trình tự có độ tương đồng ít nhất là 95%, 96%, 97%, 98%, 99%, 99,5%, 99,9%, hoặc 100% với trình tự mã hóa Yếu tố VIII-BDD được bộc lộ trong bản mô tả này, và độ tương đồng trình tự cao với phần còn lại của plasmit CS12-CRM8.2-Vrp được thể hiện trên Hình 5, ví dụ như, tương đồng ít nhất là 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99,5%, 99,9%, hoặc 100% với các nucleotit 1-528 và 4924-7794 của SEQ ID NO:10. Theo một số phương án, polynucleotit được mô tả là plasmit mà có chứa một số hoặc tất cả của các thành phần được thể hiện trong Bảng 1. Theo một số phương án, một hoặc nhiều thành phần được thể hiện trong Bảng 1 được thay thế bằng thành phần có thể so sánh được.

Sản xuất vector AAV

Polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa và vector virus được mô tả trong bản mô tả này được sản xuất theo phương pháp thông thường để khuếch đại axit nucleic và sản xuất vector. Một vài nền tảng đã được phát triển để sản xuất quy mô lớn vector AAV tái tổ hợp. Nền tảng thứ nhất dựa trên sự đưa plasmit chứa trình tự đối với hệ gen virus mong muốn vào tế bào động vật có vú chứa polynucleotit mã hóa cho các gen *rep* và *cap* AAV, cũng như là gen trợ giúp sao chép virus. Về tổng quan, xem, Kotin R.M., Hum. Mol. Genet., 20(R1):R2-6 (2011); Penaud-Budloo, M. và đồng tác giả, Mol Ther Methods Clin Dev., 8(8):166-80 (2018); và Aponte-Ubillus JJ và đồng tác giả, Appl Microbiol Biotechnol., 102(3):1045-54 (2018), nội dung của nó được kết hợp rõ ràng trong bản mô tả này để tham khảo, đến toàn bộ nội dung của nó, cho mọi mục đích. Nền tảng thứ hai dựa trên sự xây dựng dòng tế bào động vật có vú ổn định có hệ gen

virut mong muốn được tích hợp trong hệ gen tế bào động vật có vú, ví dụ như, bằng cách đồng lây nhiễm tế bào động vật có vú với adenovirut kiểu đại và adenovirut tái tổ hợp mang trình tự đối với hệ gen virut mong muốn. Về tổng quan, xem, Penaud-Budloo, M. và đồng tác giả, *Mol Ther Methods Clin Dev.*, 8(8):166-80 (2018), nội dung của nó được kết hợp rõ ràng trong bản mô tả này để tham khảo, đến toàn bộ nội dung của nó, cho mọi mục đích. Nền tảng thứ ba dựa trên sự đồng lây nhiễm của tế bào động vật có vú với HSV tái tổ hợp thứ nhất mang trình tự đối với hệ gen virut mong muốn và HSV tái tổ hợp thứ hai mã hóa cho các gen *rep* và *cap* AAV. Về tổng quan, xem, Penaud-Budloo, M. và đồng tác giả, *Mol Ther Methods Clin Dev.*, 8(8):166-80 (2018); và Aponte-Ubillus JJ và đồng tác giả, *Appl Microbiol Biotechnol.*, 102(3):1045-54 (2018), nội dung của chúng được kết hợp rõ ràng trong bản mô tả này để tham khảo, đến toàn bộ nội dung của nó, cho mọi mục đích. Nền tảng thứ tư dựa trên sự đồng lây nhiễm của tế bào côn trùng với baculovirut tái tổ hợp thứ nhất mang trình tự đối với hệ gen virut mong muốn và baculovirut tái tổ hợp thứ hai mã hóa cho các gen *rep* và *cap* AAV. Về tổng quan, xem, Penaud-Budloo, M. và đồng tác giả, *Mol Ther Methods Clin Dev.*, 8(8):166-80 (2018); và Aponte-Ubillus JJ và đồng tác giả, *Appl Microbiol Biotechnol.*, 102(3):1045-54 (2018), nội dung của chúng được kết hợp rõ ràng trong bản mô tả này để tham khảo, đến toàn bộ nội dung của nó, cho mọi mục đích. Nền tảng thứ năm dựa trên sự đưa plasmit chứa trình tự đối với hệ gen virut mong muốn vào tế bào nấm men chứa polynucleotit mã hóa cho các gen *rep* và *cap* AAV, cũng như là gen trợ giúp sao chép virut. Về tổng quan, xem, Aponte-Ubillus JJ và đồng tác giả, *Appl Microbiol Biotechnol.*, 102(3):1045-54 (2018), nội dung của nó được kết hợp rõ ràng trong bản mô tả này để tham khảo, đến toàn bộ nội dung của nó, cho mọi mục đích.

V. Phương pháp điều trị bệnh ưa chảy máu A

Theo một số phương án, hợp phần axit nucleic (ví dụ như, polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho biến thể Yếu tố VIII) và vector liệu pháp gen (ví dụ như, hạt AAV chứa polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho biến thể Yếu tố VIII) được mô tả trong bản mô tả này được dùng cho bệnh nhân mắc bệnh ưa chảy máu A để điều trị bệnh ưa chảy máu A, theo phương pháp dùng đã biết. Phương pháp dùng vector liệu pháp gen đã được biết rõ trong lĩnh vực. Chúng bao gồm, nhưng không giới hạn ở, việc dùng trong tĩnh mạch, tiêm trong cơ, tiêm kẽ, và việc dùng trong gan (ví dụ như, động mạch hoặc tĩnh mạch trong gan). Ví dụ như, xem Chuah MK và

đồng tác giả, Hum gen Ther., 23(6):557-65 (2012); Chuah MK và đồng tác giả, J Thromb Haemost., 10(8):1566-69 (2012); Chuah MK và đồng tác giả, J Thromb Haemost. 11 Suppl 1:99-110 (2013); VandenDriessche và đồng tác giả, Hum Gene Ther. 23(1):4-6 (2012); High KA, Blood, 120(23):4482-87 (2012); Matrai và đồng tác giả, Mol Ther., 18(3):477-90 (2010); và Matrai và đồng tác giả, Curr Opin Hematol., 17(5):387-92 (2010), nội dung của mỗi tài liệu này được kết hợp ở đây để tham khảo, để xem xét.

Theo đó, sáng chế đề xuất phương pháp điều trị sự thiếu hụt Yếu tố VIII (ví dụ như, bệnh ưa chảy máu A). Theo một số phương án, phương pháp này bao gồm việc dùng cho bệnh nhân cần chúng hợp phần axit nucleic (ví dụ như, polynucleotit đã được thay đổi bộ ba mã hóa cấu trúc Yếu tố VIII và/hoặc vector AAV tái tổ hợp), như được mô tả trong bản mô tả này. Theo một số phương án, hợp phần axit nucleic bao gồm polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho polypeptit biến thể Yếu tố VIII, ví dụ như, có độ tương đồng trình tự axit nucleic cao (ví dụ như, ít nhất là 95%, 96%, 97%, 98%, 99%, 99,5%, 99,9%, hoặc 100%) so với CS12-FL-NA (SEQ ID NO:1). Như được mô tả trong bản mô tả này, theo một số phương án, polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho polypeptit biến thể Yếu tố VIII được liên kết có điều khiển với vùng khởi động (ví dụ như, vùng khởi động TTR người, như được mô tả trong bản mô tả này) và một hoặc nhiều thành phần điều hòa đặc hiệu gan (ví dụ như, một hoặc hai thành phần CRM8, như được mô tả trong bản mô tả này).

Theo một số phương án, hợp phần axit nucleic là một phần của vector liệu pháp gen động vật có vú. Theo phương án cụ thể, vector liệu pháp gen động vật có vú là vector virus, ví dụ như, vector lentivirus, retrovirus, adeno virus, hoặc virus kết hợp adeno.

Theo một phương án, vector liệu pháp gen là hạt virus kết hợp adeno (AAV) mang vector virus mã hóa cho trình tự mã hóa biến thể Yếu tố VIII đã được thay đổi bộ ba mã hóa. Thông thường, vector virus bao gồm đoạn lặp đầu cuối đảo ngược (ITR) tại mỗi đầu tận cùng, một hoặc nhiều thành phần điều hòa biểu hiện (ví dụ như, vùng khởi động (ví dụ như, vùng khởi động TTR người, như được mô tả trong bản mô tả này) và một hoặc nhiều thành phần điều hòa đặc hiệu gan (ví dụ như, một hoặc hai thành phần CRM8, như được mô tả trong bản mô tả này)), trình tự mã hóa Yếu tố VIII đã được thay đổi bộ ba mã hóa, và trình tự tín hiệu poly-A.

Đánh giá hiệu quả trị liệu

Hiệu quả trị liệu của việc điều trị bệnh ưa chảy máu A có thể được đánh giá, ví dụ như, bằng cách đo tiềm năng đông máu phụ thuộc Yếu tố VIII của máu từ đối tượng được điều trị. Hệ đo để đánh giá tiềm năng đông máu bao gồm, nhưng không giới hạn ở, thử nghiệm thời gian thromboplastin một phần đã hoạt hóa (APPT) in vitro, thử nghiệm hoạt tính sinh màu Yếu tố IX, thời gian đông tụ máu, và hàm lượng kháng nguyên Yếu tố VIII (ví dụ như, bằng cách sử dụng ELISA đặc hiệu Yếu tố VIII). Cần lưu ý rằng liều lượng trị liệu cần phải không dẫn đến hàm lượng kiểu đại của Yếu tố VIII ở bệnh nhân; hơn nữa, sự biểu hiện đủ để làm giảm triệu chứng theo cách có ý nghĩa hoặc có thể đo được được coi là trị liệu đối với các mục đích của sáng chế.

Theo Tổ Chức Bệnh Ưa Chảy Máu Quốc Gia, đối tượng được phân loại là mắc bệnh ưa chảy máu A nhẹ khi huyết tương của họ chứa từ 6% đến 49% của hoạt tính Yếu tố VIII của huyết tương người bình thường. Đối tượng mắc bệnh ưa chảy máu A nhẹ thường trải qua sự chảy máu chỉ sau khi tổn thương nghiêm trọng, chấn thương hoặc phẫu thuật. Trong nhiều trường hợp, bệnh ưa chảy máu nhẹ không được chẩn đoán cho đến khi tổn thương, phẫu thuật hoặc nhổ răng dẫn đến chảy máu kéo dài. Đợt thứ nhất có thể không xảy ra cho đến tuổi trưởng thành. Phụ nữ mắc bệnh ưa chảy máu nhẹ thường trải qua rong kinh, chảy máu kinh nguyệt nặng, và có thể xuất huyết sau khi sinh con.

Theo Tổ Chức Bệnh Ưa Chảy Máu Quốc Gia, đối tượng được phân loại là có bệnh ưa chảy máu A vừa phải khi huyết tương của họ chứa từ 1% đến 5% của hoạt tính Yếu tố VIII của huyết tương người bình thường. Đối tượng mắc bệnh ưa chảy máu A vừa phải có xu hướng có các đợt chảy máu sau khi tổn thương. Sự chảy máu mà xảy ra mà không có nguyên nhân rõ ràng được gọi là các đợt chảy máu tự phát.

Theo Tổ Chức Bệnh Ưa Chảy Máu Quốc Gia, đối tượng được phân loại là có bệnh ưa chảy máu A nghiêm trọng khi huyết tương của họ chứa ít hơn 1% của hoạt tính Yếu tố VIII của huyết tương người bình thường. Đối tượng mắc bệnh ưa chảy máu A nghiêm trọng trải qua sự chảy máu sau khi tổn thương và có thể có các đợt chảy máu tự phát thường xuyên, thường vào khớp và cơ của họ.

Theo một số phương án, huyết tương người bình thường được xác định là chứa 1 IU của hoạt tính Yếu tố VIII trong mỗi ml. Do đó, theo một số phương án, huyết tương

từ đối tượng được phân loại là mắc bệnh ưa chảy máu A nhẹ chứa từ 0,06 và 0,49 IU của hoạt tính Yếu tố VIII trong mỗi ml. Theo một số phương án, huyết tương từ đối tượng được phân loại là mắc bệnh ưa chảy máu A vừa phải chứa từ 0,01 đến 0,05 IU của hoạt tính Yếu tố VIII trong mỗi ml. Theo một số phương án, huyết tương từ đối tượng được phân loại là mắc bệnh ưa chảy máu A nặng chứa từ 0,01 đến 0,05 IU của hoạt tính Yếu tố VIII trong mỗi ml.

Theo một số phương án, liệu pháp là hữu hiệu để điều trị khi nó làm giảm độ nghiêm trọng của triệu chứng của bệnh ưa chảy máu A, ví dụ như, bằng cách làm tăng mức độ trung bình của hoạt tính Yếu tố VIII trong máu của đối tượng. Theo đó, theo một số phương án, liệu pháp bệnh ưa chảy máu A hữu hiệu để điều trị khi nó làm tăng mức độ trung bình của hoạt tính Yếu tố VIII ở máu/huyết tương của đối tượng. Theo một số phương án, việc điều trị có tác dụng trị liệu làm tăng mức độ trung bình của hoạt tính Yếu tố VIII ở máu/huyết tương của đối tượng bằng ít nhất là 0,5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, hoặc lớn hơn.

Theo một số phương án, việc điều trị hữu hiệu về mặt trị liệu làm tăng mức độ trung bình của hoạt tính Yếu tố VIII ở máu của đối tượng sao cho đối tượng được phân loại là có dạng ít nghiêm trọng hơn của bệnh ưa chảy máu A. Ví dụ như, trong một phương án, đối tượng ban đầu được phân loại là mắc bệnh ưa chảy máu A nặng được phân loại lại là mắc bệnh ưa chảy máu A vừa phải hoặc bệnh ưa chảy máu A nhẹ sau khi trải qua việc điều trị hữu hiệu về mặt trị liệu. Theo phương án khác, đối tượng ban đầu được phân loại là mắc bệnh ưa chảy máu A vừa phải được phân loại lại là mắc bệnh ưa chảy máu A nhẹ sau khi trải qua việc điều trị hữu hiệu về mặt trị liệu. Theo phương án khác, đối tượng ban đầu được phân loại là mắc bệnh ưa chảy máu A nhẹ được phân loại lại là không mắc bệnh ưa chảy máu A sau khi trải qua việc điều trị hữu hiệu về mặt trị liệu.

Chế phẩm

Hợp phân để sử dụng trong việc điều trị bệnh ưa chảy máu A được đề xuất trong bản mô tả này. Hợp phân này chứa lượng hữu hiệu để điều trị của hợp phân axit nucleic, ví dụ như, vectơ liệu pháp gen AAV bao gồm polynucleotit đã được thay đổi bộ ba mã hóa mã hóa cho Yếu tố VIII, như được mô tả trong bản mô tả này. Lượng hữu hiệu để

điều trị của polynucleotit VIII đã được thay đổi bộ ba mã hóa (ví dụ như, vector liệu pháp gen AAV bao gồm trình tự mã hóa Yếu tố VIII đã được thay đổi bộ ba mã hóa) được trộn với chất mang hoặc tá dược lỏng dùng cho được phẩm thích hợp để, ví dụ như, dùng toàn thân. Việc tạo chế phẩm cuối cùng của polynucleotit Yếu tố VIII đã được thay đổi bộ ba mã hóa được bộc lộ trong bản mô tả này nằm trong khả năng của người có trình độ trung bình trong lĩnh vực.

Liều lượng

Hợp phần axit nucleic theo sáng chế được dùng cho bệnh nhân cần chúng. Lượng hoặc liều lượng của tác nhân liệu pháp gen trị liệu được dùng phụ thuộc vào các yếu tố chẳng hạn như cấu trúc polynucleotit VIII đã được thay đổi bộ ba mã hóa cụ thể, vector phân phối được sử dụng, mức độ nghiêm trọng của bệnh, và đặc điểm chung của đối tượng. Liều lượng chính xác sẽ phụ thuộc vào mục đích điều trị, và sẽ được xác định bởi người có hiểu biết trong lĩnh vực kỹ thuật bằng cách sử dụng các kỹ thuật đã biết (xem, ví dụ như, Lieberman, *Pharmaceutical Dosage Forms* (vols. 1-3, 1992); Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999); Pickar, *Dosage Calculations* (1999); và Remington: *The Science and Practice of Pharmacy*, ấn bản thứ 20, 2003, Gennaro, Ed., Lippincott, Williams & Wilkins). Người có trình độ trung bình trong lĩnh vực có khả năng xác định liều lượng và phác đồ dùng liều lượng cụ thể để điều trị đối tượng cụ thể.

Ví dụ thực hiện sáng chế

Ví dụ 1 - Cải thiện cấu trúc vector AAV8 sợi đơn biểu hiện FVIII bằng cách bổ sung các trình tự cầu nối N-glycosyl hóa

Để biết được liệu các vị trí N-glycosyl hóa bổ sung ở trong trình tự SQ của BDD-FVIII có làm tăng sự biểu hiện protein FVIII hay không, tập hợp của các trình tự peptit ngắn chứa các vị trí glycosyl hóa liên kết N giả định được thiết kế. Trước đây, McIntosh và cộng sự (*Blood* 121(17):3335-44 (2013)) đã chứng minh rằng ý tưởng về sự N-glycosyl hóa với 6 vị trí tiềm năng (“peptit V3”) dẫn đến mức độ biểu hiện FVIII được tăng cường trong huyết tương của chuột nhắt. Điều thú vị là, sự dự đoán in silico của “peptit V3” xác định hai trong số 6 vị trí có khả năng được N-glycosyl hóa in vivo.

12 trình tự cầu nối khác nhau, được thể hiện dưới dạng NG1-NG21 trên các Hình 8A-8B, được thiết kế ở trong ngữ cảnh của trình tự BDD-FVIII “CS01” được tối ưu hóa bộ ba mã hóa (xem, WO 2017/083762, nội dung của nó được kết hợp ở đây trong bản mô tả này để tham khảo) bằng cách áp dụng chương trình NetNGlyc-Database (Steentoft và đồng tác giả, H. EMBO J, 32(10):1478-88, 2013), với mục đích tạo ra các trình tự ngắn chứa nhiều vị trí N-glycosyl hóa. Nền tảng NetNGly kết hợp chín mạng lưới thần kinh phân tích các trình tự protein người về kiểu N-glycosyl hóa của chúng. Dựa trên cơ sở dữ liệu NetNGly, peptit đã được thiết kế được phân tích về khả năng xảy ra để truyền sự N-glycosyl hóa dưới dạng sự cải biến sau dịch mã, như được mô tả trong Steentoft và đồng tác giả, H. EMBO J, 32(10):1478-88, (2013), nội dung của nó được kết hợp ở đây để tham khảo. Trong số 12 peptit thay thế mới, bốn cầu nối NG hứa hẹn với khả năng gây miễn dịch thấp được dự đoán (như được giải thích dưới đây) được cài xen vào trong trình tự SQ có kích thước 14 axit amin (SFSQN – peptit mới – PPVLKRHRQ) của trình tự BDD-FVIII được tối ưu hóa bộ ba mã hóa được gọi là “CS04” (SEQ ID NO:10).

Sự cải biến sau dịch mã của vị trí N-glycosyl hóa đã được dự đoán được xác nhận bằng thí nghiệm đối với ba vector bao gồm vNG4/CS04, vNG5/CS04, và vNG16/CS04, bằng cách chuyển nạp tế bào Huh-7 gan người và phát hiện BDD-FVIII đã được cải biến bằng phân tích thẩm tách Western kháng-FVIII (Hình 9). Khi so với BDD-FVIII chỉ chứa trình tự SQ, chuỗi nặng có kích thước lớn của các phiên bản đã được cải biến của FVIII được phát hiện bằng cách điện di trên gel, chỉ ra sự có mặt của vị trí N-glycosyl hóa mới. Ngoài ra, sự lây nhiễm AAV8 của dòng tế bào gan người HepG2 xác nhận sự N-glycosyl hóa như được thể hiện ví dụ đối với vector biểu hiện vNG5/CS04 và biến thể X5 tiết cao vX5/NG5/CS125 (Hình 10; vX5/NG5/CS125 được mô tả trong Ví dụ 3 và Ví dụ 5).

Sự tập trung vào thiết kế trình tự cầu nối là để tạo ra các trình tự peptit ngắn (7-17 gốc axit amin) có nguy cơ sinh miễn dịch thấp. Vì điều này, việc lập biên dạng khả năng sinh miễn dịch in silico Epibase™ (Lonza) được áp dụng (sự liên kết HLA). Phương pháp dự đoán epitop Epibase™ có chứa lớp cấu trúc và thống kê để dự đoán khả năng sinh miễn dịch. Phần cấu trúc ước tính ái lực liên kết dựa trên công nghệ Pepscope, như được mô tả trong Desmet và đồng tác giả, Proteins. 58(1):53-69 (2005), nội dung của nó được kết hợp ở đây trong bản mô tả này để tham khảo. Phần thống kê rút ra thông

tin từ các trình tự peptit và ái lực liên kết thí nghiệm của chúng. Dựa trên số đếm epitop tới hạn, các alotyp HLA lớp II bị ảnh hưởng (Krischmann và đồng tác giả, J Immunol. 15;155(12):5655-62 (1995); Verreck và đồng tác giả, Int Immunol. 8(3):397-404 (1996), nội dung của nó được kết hợp ở đây trong bản mô tả này để tham khảo) và điểm nguy cơ DRB1 (Laupeze và đồng tác giả, Hum Immunol. 60(7):591-7 (1999), nội dung của nó được kết hợp ở đây trong bản mô tả này để tham khảo) được sử dụng để ước tính nguy cơ gây miễn dịch của các protein. Dựa trên hệ thống tính điểm này, nguy cơ gây miễn dịch của peptit NG16 là vừa phải, nguy cơ gây miễn dịch của NG4 và NG10 là thấp và của NG5 thậm chí thấp hơn trình tự SQ không được cải biến của BDD-FVIII (Bảng 2).

Bảng 2. Hoạt tính FVIII hai tuần sau khi điều trị AAV *in vivo*.

Cấu trúc AAV	Cầu nối			Epibase	Chuột nhắt bị knock-out FVIII		Chuột nhắt “dòng E”	
	ID	Số lượng gốc axit amin	Số lượng vị trí N-glycosyl hóa đã được dự đoán		Hoạt tính FVIII [IU/ml]	Số lần tăng lên so với Orth 04	Hoạt tính FVIII [IU/ml]	Số lần tăng lên so với Orth04
vCS04	-	0	0	thấp	$1,7 \pm 0,5$	1,0	$1,4 \pm 0,6$	1,0
vNG4/CS04	NG4	11	3	thấp	$3,1 \pm 0,6$	1,8	không phát hiện	không phát hiện
vNG5/CS04	NG5	7	2	rất thấp	$3,8 \pm 0,8$	2,3	$4,0 \pm 2,8$	3,0
vNG10/CS04	NG10	17	2	thấp	$4,1 \pm 0,5$	2,4	không phát hiện	không phát hiện
vNG16/CS04	NG16	9	2	vừa phải	$2,5 \pm 0,8$	1,4	không phát hiện	không phát hiện

Hơn nữa, vector vNG4/CS04, vNG5/CS04, vNG10/CS04 và vNG16/CS04 được phân tích trong mô hình chuột nhắt bị knock-out exon 16 FVIII được tạo ra trong phòng thí nghiệm của Haig Kazazian tại University of Pennsylvania (Bảng 2). Hai tuần sau khi lây nhiễm AAV8, mức độ biểu hiện FVIII được xác định bằng hoạt tính sinh màu của mẫu huyết tương chuột nhắt. Hai tuần sau khi lây nhiễm AAV8 ở liều lượng bằng

$4,0 \times 10^{12}$, mức độ biểu hiện FVIII cao hơn từ 1,4 đến 2,4 lần so với cấu trúc vCS04 không chứa peptit. Nhìn chung, cấu trúc vNG5/CS04 thể hiện các đặc điểm phù hợp nhất, bao gồm mức độ biểu hiện FVIII cao (3,8 IU/ml), nguy cơ gây miễn dịch thấp nhất, và peptit có kích thước ngắn nhất (7 gốc axit amin). Cấu trúc vNG5/CS04 được thử nghiệm thêm trong mô hình chuột nhắt FVIII khác, mô hình “dòng E”, mà dung chịu miễn dịch đối với FVIII người (được mô tả trong Reipert và đồng tác giả, *Haemophilia*. 16 (Suppl 5):47-53 (2010) và van Helden và đồng tác giả, *Blood* 118(13):3698-707 (2011), nội dung của nó được kết hợp rõ ràng trong bản mô tả này để tham khảo). Trong mô hình chuột nhắt độc lập này, sự biểu hiện FVIII tăng lên của vNG5/CS04 được xác nhận và được đánh giá là cao hơn 3 lần so với vCS04.

Cùng nhau, vector được cải thiện được gọi là vNG5/CS04 chứa trình tự peptit N-glycosyl hóa mới ngắn với nguy cơ gây miễn dịch rất thấp và mức độ được tăng cường của sự biểu hiện FVIII in vivo được phát triển.

Ví dụ 2 - Cải thiện cấu trúc vector AAV8 sợi đơn biểu hiện FVIII bằng cách thay thế vùng khởi động/vùng tăng cường

Cát xét vùng khởi động của cấu trúc vector sợi đơn dựa trên AAV8 biểu hiện FVIII ban đầu được gọi là vCS04 chứa trình tự vùng khởi động và vùng tăng cường lõi bắt nguồn từ gen transthyretin (TTR) chuột đặc hiệu gan (Yan và đồng tác giả, *EMBO J*, 9:869-78 (1990), nội dung của nó được kết hợp ở đây để tham khảo), xem Hình 11. Cát xét vùng khởi động/vùng tăng cường mTTR đặc hiệu gan này được cải biến hai lần, trước hết trình tự vùng khởi động lõi (nền) của chuột được thay thế đầy đủ bằng trình tự của người tương ứng. Thứ hai, trình tự vùng tăng cường của chuột được thay thế bằng mô đun điều hòa cix đặc hiệu gan được mô tả gần đây CRM8 (Nair và đồng tác giả, *Blood*, 123:3195-99 (2014) và Chuah và đồng tác giả, *Mol Ther*, 22:1605-13 (2014), nội dung của nó được kết hợp ở đây trong bản mô tả này để tham khảo). Một hoặc hai thành phần CRM8 được cài xen ngược dòng của vùng khởi động lõi (nền) TTR của người, lần lượt dẫn đến cấu trúc vector vCS115 và vCS116 (Hình 11).

Độ mạnh của cát xét vùng khởi động mới được đánh giá bằng các phân tích in vitro và in vivo, trong dòng tế bào HepG2 có nguồn gốc từ gan người và ở chuột nhắt “dòng E” ở liều lượng bằng $4,0 \times 10^{12}$ vg/kg. In vitro, các cấu trúc có một thành phần CRM8 (vCS115) và hai thành phần CRM8 (vCS116) dẫn đến đơn vị hiệu lực sinh học

cao hơn lần lượt 2,2 và 3,7 lần, khi so với tham chiếu (vCS04) (Hình 12). In vivo, cát xét vùng khởi động CRM8/hTTR được thực hiện có thể so sánh được với vùng khởi động/vùng tăng cường mTTR (Hình 12).

Tác dụng gia tăng liên quan đến CRM8 trên sự biểu hiện FVIII được đánh giá thêm trong các tập hợp ba thành phần của các cấu trúc, trong đó các cát xét vùng khởi động được mô tả bao gồm vùng khởi động/vùng tăng cường mTTR, 1xCRM8/hTTR, và 2xCRM8/hTTR được kết hợp với sự cải biến mới của BDD-FVIII (xem các Ví dụ 3 -5 dưới đây). Nhìn chung, dữ liệu in vitro cho thấy sự tăng lên của đơn vị hiệu lực sinh học dựa trên sự có mặt của một và hai thành phần CRM8 bằng cách so sánh trực tiếp (1) vNG5/CS04 với vNG5/CS117 (cao hơn 2,5 lần) và vNG5/CS118 (cao hơn 4,0 lần), (2) vX5/CS24 với vX5/CS101 (cao hơn 1,8 lần) và vX5/CS105 (cao hơn 5,3 lần) và (3) vX5/NG5/CS125 với vX5/NG5/CS119 (cao hơn 4,4 lần) và vX5/NG5/CS120 (cao hơn 6,2 lần). Không có tác dụng phụ thuộc CRM8 lên hàm lượng FVIII được quan sát thấy in vivo. Tuy nhiên, dữ liệu in vivo cho thấy rõ ràng rằng cả hai nền tảng cát xét vùng khởi động 1xCRM8/hTTR và 2xCRM8/hTTR mới tốt ngang nhau trong mô hình chuột nhất được so sánh với vùng khởi động/vùng tăng cường mTTR.

Ví dụ 3 - Cải thiện cấu trúc vector AAV8 sợi đơn biểu hiện FVIII bằng cách đưa biến thể “NG5” vào

Dựa trên dữ liệu được thể hiện trong Ví dụ 1, cầu nối N-glycosyl hóa NG5 trong cấu trúc vNG5/CS04 được chọn để kết hợp với vùng khởi động đặc hiệu gan người mới 1xCRM8/hTTR và 2xCRM8/hTTR, dẫn đến cấu trúc vNG5/CS117 và vNG5/CS118 (Hình 11). Cấu trúc vNG5/CS04 chứa cát xét vùng khởi động/vùng tăng cường mTTR (Hình 11). Mức độ hiệu lực sinh học in vivo của vNG5/CS04 cao hơn 1,9 lần so với vCS04, do tác dụng dương tính trên trình tự cầu nối NG5 (Hình 12). In vitro, sự tăng lên của hiệu lực sinh học được quan sát thấy đối với vector chứa (các) thành phần CRM8 vNG5/CS117 và vNG5/CS118. In vivo, mức độ biểu hiện của vNG5/CS117 và vNG5/CS118 tương tự với vNG5/CS04 (Hình 12).

Ví dụ 4 - Cải thiện cấu trúc vector AAV8 sợi đơn biểu hiện FVIII bằng cách đưa biến thể “X5” vào

Để cải thiện hơn nữa vector biểu hiện FVIII, biến thể X5 (được mô tả dưới dạng đột biến ‘m2’ trong WO 2017/083762, nội dung của nó được kết hợp ở đây trong bản

mô tả này để tham khảo) được đưa vào BDD-FVIII. Biến thể X5 chứa năm gốc axit amin FVIII lộn ở trong miền A1 của chuỗi nặng mà mang lại sự tiết hiệu quả cho FVIII người (Cao và đồng tác giả, 2014, tóm tắt ASGCT #460; chi tiết về các biến thể được bộc lộ trong cách trình bày dùng qua đường miệng). Cụ thể là, biến thể X5 của BDD-FVIII được kết hợp với vùng khởi động/vùng tăng cường mTTR, dẫn đến vX5/CS24, và với một hoặc hai thành phần CRM8 cộng với vùng khởi động hTTR, dẫn đến cấu trúc vX5/CS101 và vX5/CS105 (Hình 11). Ba cấu trúc mới được phân tích in vitro trong dòng tế bào HepG2 có nguồn gốc từ gan và in vivo trong chuột nhắt “dòng E2”, dòng chuột nhắt bị knock-out FVIII chuyển gen mà biểu hiện lượng tối thiểu của cADN FVIII người dẫn đến sự dung chịu miễn dịch đối với FVIII người (van Helden PM, và đồng tác giả, Blood 118(13):3698-707 (2011), nội dung của nó được kết hợp để tham khảo trong bản mô tả này, đến toàn bộ nội dung của nó, cho mọi mục đích), và được so sánh với cấu trúc tham chiếu vCS04. In vivo, vX5/CS24, vX5/CS101, và vX5/CS105 dẫn đến mức độ được tăng cường của hoạt tính FVIII theo hệ số bằng 2,5 – 3,1 (Hình 12). In vitro, sự tăng lên của hiệu lực sinh học nằm trong khoảng từ 4,0 đến 21,2. Hệ thống thử nghiệm in vitro phát hiện ra sự twang lên 4,0 lần nảy sinh từ biến thể X5 (cấu trúc vX5/CS24), và được làm tăng thêm theo hệ số 7,3 và 21,2 bằng các cấu trúc chứa X5 và một hoặc hai bản sao của thành phần CRM8 (lần lượt cấu trúc vX5/CS101 và vX5/CS105).

Ví dụ 5 - Cải thiện cấu trúc vector AAV8 sợi đơn biểu hiện FVIII bằng cách đưa cả hai biến thể “X5” và “NG5” vào

Trong tập hợp khác của vector, cầu nối N-glycosyl hóa NG5 (Ví dụ 3) và biến thể X5 (Ví dụ 4) được đưa vào BDD-FVIII song song, và kết hợp thêm với ba vùng khởi động khác nhau, bao gồm vùng khởi động/vùng tăng cường mTTR, 1x CRM8/hTTR, và 2xCRM8/hTTR, lần lượt dẫn đến cấu trúc vX5/NG5/CS125, vX5/NG5/CS119 và vX5/NG5/CS120 (Hình 11). Khi so với vCS04, hiệu lực sinh học in vivo và in vitro của vX5/NG5/CS125, vX5/NG5/CS119, và vX5/NG5/CS120 được đánh giá lần lượt theo hệ số 3,6 – 5,5 và 3,2 – 19,8. Sự đưa vào của hai cải biến mới X5 và NG5 phát hiện thêm sự biểu hiện tăng, in vivo, như được thể hiện đối với ba loạt cấu trúc mỗi loạt có vùng khởi động chung. Ví dụ như, đối với các loạt có vùng khởi động 2xCRM8/hTTR (cấu trúc vCS116, vNG5/CS118, vX5/CS105, và vX5/NG5/CS120), sự có mặt của một mình NG5 làm tăng mức độ biểu hiện FVIII từ 1,9 đến 2,9 IU/ml, sự có mặt của một

mình X5 làm tăng sự biểu hiện FVIII từ 1,9 đến 5,1 IU/ml, và sự có mặt của cả X5 và NG5 làm tăng sự biểu hiện FVIII từ 1,9 đến 11,4 IU/ml.

Cùng nhau, cấu trúc vX5/NG5/CS120 mang cát xét vùng khởi động 2xCRM8/hTTR mới và thêm trình tự nucleotit BDD-FVIII “CS04” được tối ưu hóa bộ ba mã hóa với hai cải biến, biến thể X5 và trình tự NG5, được xác định là cấu trúc có hiệu lực sinh học in vitro và in vivo cao nhất.

Ví dụ 6 - Sự giảm nucleotit của cấu trúc vector AAV8 sợi đơn

Để làm giảm kích thước vector của cấu trúc vector hơi quá khổ vX5/NG5/CS120 mà dẫn đến sự đóng gói hiệu quả hơn, tất cả các trình tự ADN không hoạt động chức năng ở trong ITR kẹp hai bên được làm khuyết. Điều này dẫn đến sự giảm đi của tổng cộng 71 nucleotit, từ cát xét biểu hiện có kích thước 5191 nucleotit thành cát xét biểu hiện có kích thước 5120 nucleotit. Sự làm khuyết bao gồm 19 nucleotit giữa 5'-ITR và trình tự CRM8, 9 nucleotit giữa vùng khởi động TTR người và trình tự Kozak, 27 nucleotit giữa trình tự mã hóa BDD-FVIII và vị trí polyadenyl hóa tổng hợp, và 16 nucleotit giữa vị trí polyadenyl hóa tổng hợp và trình tự 3'-ITR. Cát xét biểu hiện đã được giảm kích thước, được gọi là vX5/NG5/CS12, gồm có vùng khởi động với hai thành phần CRM8 và trình tự vùng khởi động TTR người lõi, trình tự BDD-FVIII bao gồm biến thể X5 cũng như là trình tự NG5 và vị trí polyadenyl hóa tổng hợp và được kẹp hai bên bằng hai đoạn lặp đầu cuối đảo ngược dựa trên AAV2. Cấu trúc vX5/NG5/CS12 được thể hiện bằng sơ đồ trên Hình 11.

Tính toán vẹn hệ gen vector của các chế phẩm hệ gen vector AAV vCS04, vX5/NG5/CS120 và vX5/NG5/CS12 được nhắm đến bằng điện di trên gel agarosa. Kết quả được thể hiện trên Hình 13 chứng tỏ rằng các vector virus vCS04, vX5/NG5/CS120 và vX5/NG5/CS12 có hệ gen có kích thước tương tự, được chỉ ra bởi băng phác biệt xấp xỉ 5kb. Dù kích thước vector tính được bằng xấp xỉ 5,2 kb, hệ gen là băng đồng nhất xác nhận việc đóng gói đúng của hệ gen quá cỡ ở rìa (tương quan với hệ gen kiểu đại AAV 4,7 kb). Biến thể vX5/NG5/CS12 ngắn hơn được ưu tiên.

Tập hợp của vector virus dựa trên AAV8 bao gồm vCS04, vX5/NG5/CS120, và vX5/NG5/CS12 được dùng cho chuột nhắt bị knock-in FVIII F17 ở liều lượng vector bằng 5×10^{11} vg/kg, 1×10^{12} vg/kg, và 4×10^{12} vg/kg và mức độ hoạt tính FVIII được xác định vào ngày 14. Như được thể hiện trên Hình 14, cả vX5/NG5/CS12 và

vX5/NG5/CS120 dẫn đến mức độ biểu hiện có thể so sánh được bằng xấp xỉ 4 IU/ml ở liều lượng vector bằng 1×10^{12} vg/kg. Trái ngược với cấu trúc tham chiếu vCS04 có mức độ biểu hiện FVIII rất thấp bằng 0,3 IU/ml ở liều lượng bằng 1×10^{12} vg/kg, các vector đã được cải thiện vX5/NG5/CS12 và vX5/NG5/CS120 thể hiện mức độ biểu hiện FVIII tăng lên theo hệ số bằng xấp xỉ 14 (Hình 14). Thậm chí ở liều lượng bằng 5×10^{11} vg/kg mức độ biểu hiện của 1,5 IU/ml FVIII có thể thu được đối với vector vX5/NG5/CS12. Theo dữ liệu in vivo, hiệu lực sinh học in vitro của vX5/NG5/CS120 và vX5/NG5/CS12 được tăng mạnh xấp xỉ theo hệ số lần lượt là 17 và 24, lây nhiễm tế bào HepG2 ở bội số lây nhiễm (MOI) ngang bằng (Hình 14). Do đó, bởi vì tất cả vector được dùng ở cùng nồng độ (bội số lây nhiễm), sự khác nhau ở hiệu lực sinh học là do các biến thể được sử dụng và không phụ thuộc vào kích thước.

YÊU CẦU BẢO HỘ

1. Hợp phần axit nucleic có chứa polynucleotit Yếu tố VIII mã hóa cho protein Yếu tố VIII, polynucleotit Yếu tố VIII này có trình tự axit nucleic của SEQ ID NO:1.
2. Hợp phần axit nucleic theo điểm 1, còn chứa vùng khởi động được liên kết có điều khiển với polynucleotit Yếu tố VIII, trong đó polynucleotit vùng khởi động có trình tự axit nucleic của SEQ ID NO:6.
3. Hợp phần axit nucleic theo điểm 2, trong đó không có nucleotit giữa vùng khởi động và polynucleotit Yếu tố VIII.
4. Hợp phần axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 3, còn chứa thành phần đặc hiệu gan được liên kết có điều khiển với polynucleotit Yếu tố VIII, trong đó thành phần đặc hiệu gan có trình tự của SEQ ID NO:5.
5. Hợp phần axit nucleic theo điểm 4 trong đó thành phần đặc hiệu gan thứ hai được liên kết có điều khiển với polynucleotit Yếu tố VIII.
6. Hợp phần axit nucleic theo điểm 4 hoặc 5, còn chứa vùng khởi động được liên kết có điều khiển với polynucleotit Yếu tố VIII, trong đó:

thành phần đặc hiệu gan ở vị trí ngược dòng của vùng khởi động,

vùng khởi động được bố trí phía trước polynucleotit yếu tố VIII, và

không có nucleotit nằm giữa thành phần đặc hiệu gan và vùng khởi động.
7. Hợp phần axit nucleic theo điểm 1, có trình tự axit nucleic có chứa SEQ ID NO: 3.
8. Vector liệu pháp gen động vật có vú có chứa hợp phần axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 7.
9. Vector liệu pháp gen động vật có vú theo điểm 8, trong đó vector liệu pháp gen động vật có vú là vector virus kết hợp adeno (adeno-associated virus - AAV).
10. Vector liệu pháp gen động vật có vú theo điểm 9, trong đó vector AAV là vector virus kết hợp adeno tít huyết thanh 8 (AAV-8).
11. Vector liệu pháp gen động vật có vú theo điểm bất kỳ trong số các điểm từ 8 đến 10, trong đó polynucleotit Yếu tố VIII mã hóa cho protein Yếu tố VIII là polynucleotit sợi đơn có trình tự axit nucleic của SEQ ID NO: 1.

12. Hạt virus kết hợp adeno (AAV) có chứa protein capsit bao nang hợp phần axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 7.

13. Hạt AAV theo điểm 12, trong đó protein capsit có chứa protein capsit virus kết hợp adeno tít huyết thanh 8 (AAV-8).

14. Hạt AAV theo điểm 12 hoặc 13, trong đó polynucleotit Yếu tố VII mã hóa cho protein Yếu tố VIII là polynucleotit sợi đơn có trình tự axit nucleic của SEQ ID NO: 1.

15. Phương pháp sản xuất hạt virus kết hợp adeno (AAV) có chứa việc đưa hợp phần axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 7 vào tế bào chủ nhân chuẩn: trong đó:

hợp phần axit nucleic có chứa polynucleotit Yếu tố VIII có trình tự axit nucleic của SEQ ID NO:1 được kẹp hai bên bằng trình tự đoạn lặp đầu cuối đảo ngược 5' (5' ITR) và trình tự đoạn lặp đầu cuối đảo ngược 3' (3' ITR), và

tế bào chủ nhân chuẩn có chứa một hoặc nhiều polynucleotit mã hóa cho gen *rep* AAV, gen *cap* AAV, và gen trợ giúp sao chép virus.

16. Phương pháp theo điểm 15, trong đó hợp phần axit nucleic là plasmit có trình tự axit nucleic của SEQ ID NO:10.

17. Phương pháp theo điểm 15 hoặc 16, trong đó gen *cap* AAV là gen *cap* virus kết hợp adeno tít huyết thanh 8 (AAV-8).

1/28

CS12-FL-NA

atgcagattgagctgagcacctgcttcttctgtgcctgctgaggttctgcttcttctgccaccagga
gatactacctgggggctgtggagcttcttgggactacatgcagctctgacctgggggagctgacctgt
ggatgccagggttcccacccagagtgccaaatccttccattcaacacctctgtggtctacaagaag
acctcttctgtggagttcactgaccacctgttcaacattgccaaacccaggccacctggatgggac
tcttgggacccaccattcaggctgaggtgtatgacactgtggtcgtcacctcaagaacatggcctc
ccacctgtgagcctgcatgctgtgggggtcagctactggaagtcctctgagggggctgagtatgat
gaccagacctcccagagggagaaggaggatgacaaagtgttccctgggaagagccacacctatgtgt
ggcagggtcctcaaggagaatggccccactgcctctgacccacctgcctgacctactcctaccttct
tcatgtggacctggtcaaggacctcaactctggactgattggggccctgctggtgtgagggagggc
tccctggccaaagagaagaccagacctgcacaagttcattctcctgtttgctgtctttgatgagg
gcaagagctggcactctgaaaccaagaactcctgatgcaggacagggatgctgcctctgccagggc
ctggcccaagatgcacactgtgaatggctatgtgaacaggagcctgcctggactcattggctgccac
aggaaatctgtctactggcatgtgattggcatggggacaacccctgaggtgcactccatttctctgg
agggccacaccttctggtcaggaaccacagacagggccagcctggagatcagccccatcaccttctc
cactgccagacctgctgatggacctcggacagttcctgctgttctgccacatcagctcccaccag
catgatggcatggaggcctatgtcaaggtggacagctgccctgaggagccacagctcaggatgaaga
acaatgaggaggctgaggactatgatgatgacctgactgactctgagatggatgtggtccgctttga
tgatgacaacagcccatccttcattcagatcaggtctgtggccaagaaacaccccaagacctgggtg
cactacattgctgctgaggaggaggactgggactatgccccactggtcctggccccctgatgacagga
gctacaagagccagctacctcaacaatggccccacagaggattggacgcaagtacaagaaagtcaggtt
catggcctacactgatgaaaccttcaagaccagggaggccattcagcatgagctctggcatcctgggc
ccactcctgtatggggaggtgggggacacctgctcatcatcttcaagaaccaggcctccaggccct
acaacatctaccacatggcatcactgatgtcaggccccctgtacagccgcaggctgcccagggggt
gaaacacctcaaggacttccccattctgctggggagatcttcaagtacaagtggactgtcactgtg
gaggatggaccaaccaaatctgacccaggtgcctcaccagatactactccagctttgtgaacatgg
agagggacctggcctctggcctgattggccccactgctcatctgctacaaggagtctgtggaccagag
gggaaaccagatcatgtctgacaagaggaatgtgattctgttctctgtctttgatgagaacagggac
tggtacctgactgagaacattcagcgttctgccccacccctgctggggtgcagctggaggacctg
agttccaggccagcaacatcatgcactccatcaatggctatgtgtttgacagcctccagctttctgt
ctgctgcatgaggtggcctactgggtacattcttcttattggggcccagactgacttcttctgtc
ttcttctctggctacaccttcaaacacaagatggtgtatgaggacacctgacctcttccattct
ctgggggagactgtgttcatgagcatggagaacctggcctgtggattctgggatgccacaactctga
cttccgcaacaggggcatgactgcctgctcaaagtctcctcctgtgacaagaacactggggactac
tatgaggacagctatgaggacatctctgcctacctgctcagcaagaacaatgccattgagcccagga
gcttcagccagaatgtgagcaataatgccaccaacccacctgtcctgaaacgccaccagagggagat
caccaggaccacctccagctctgaccaggaggagattgactatgatgacaccatttctgtggagatg
aagaaagaggactttgacatctatgacgaggacgagaaccagagcccaaggagcttcagaagaaga
ccaggcaCtacttcattgctgctgtggagcgctgtgggactatggcatgagctccagccccatgt
cctcaggaacagggccagctctggctctgtgccacagttcaagaaagtggctcttccaagagttcact

(Tiếp)

Hình 1A

2/28

gatggcagcttcacccagcccctgtacagaggggagctgaatgagcacctgggactcctgggcccac
acatcagggctgaggtggaggacaacatcatggtgaccttcgcgaaccaggcctccaggccctacag
cttctacagctccctcatcagctatgaggaggaccagaggcagggggctgagccacgcaagaacttt
gtgaaacccaatgaaaccaagacctacttcttgaaagtccagcaccacatggccccaccaaggatg
agtttgactgcaaggcctgggcctacttctctgatgtggacctggagaaggatgtgcactctggcct
gattggcccactcctggtctgccacaccaacacctgaacctgcccattggaaggcaagtactgtg
caggagtttgccctcttcttcaccatctttgatgaaaccaagagctggtacttcactgagaacatgg
agcgcaactgcagggccccatgcaacattcagatggaggacccccaccttcaaagagaactaccgctt
ccatgccatcaatggctacatcatggacacctgacctgggcttgcatggcccaggaccagaggatc
aggtggtacctgctttctatgggctccaatgagaacattcactccatccacttctctgggcatgtct
tcaactgtgcgcaagaaggaggagtacaagatggccctgtacaacctctacctggggtctttgagac
tgtggagatgctgcccctcaaagctggcatctggagggaggagtgcctcattggggagcacctgcat
gctggcatgagcacctgttctctggtctacagcaacaagtgccagacccccctgggaatggcctctg
gccacatcagggacttccagatcactgcctctggccagatggccagtgggcccccaagctggccag
gctccactactctggatccatcaatgcctggagcaccaaggagccattcagctggatcaaagtggac
ctgctggcccccatgatcatccatggcatcaagacccagggggccaggcagaagttctccagcctgt
acatcagccagttcatcatcatgtacagcctggatggcaagaaatggcagacctacagaggcaactc
cactggaacactcatggtcttctttggcaatgtggacagctctggcatcaagcacaacatcttcaac
cccccaatcatcgccagatacatcaggctgcacccccaccactacagcatccgcagcacccctcagga
tgagctgatgggctgtgacctgaactcctgcagcatgcccctgggcatggagagcaaggccatttc
tgatgccagatcactgcctccagctacttcaccaacatgtttgccacctggagcccgaaggcc
aggctgcacctccagggaaggagcaatgcctggaggccccagggtcaacaacccaaaggagtggctgc
aggtggacttccagaagaccatgaaggctcactggggtgaccacccaggggtcaagagcctgctcac
cagcatgtatgtgaaggagtctcctgatcagctccagccaggatggccaccagtggaacctcttcttc
cagaatggcaagggtcaagggtgttccagggcaaccaggacagcttCacccctgtggtgaacagcctgg
acccccccctcctgaccagatacctgaggattcacccccagagctgggtccaccagattgccctgag
gatggaggtcctgggatgtgaggcccaggacctgtac (SEQ ID NO:1)

Hình 1B

3/28

CS12-FL-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPPN
TSVVYKKTLEFVEFTDHLFNIAPRPPWMGLLGPTIQAEVYDTVVVTLKNMASHPVSLHAV
GVSYWKSSEGAEYDDQTSQREKEDDKVFPKSHTYVWQVLKENGPTASDPPCLTYSYLSH
VDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNLSMQDRD
AASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNH
RQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNE
EAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLA
PDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGDTL
LIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGP
TKSDPRCLTRYYSFVNMERDLASGLIGPLLYCYKESVDQRGNQIMSDKRNVLFSVFDE
NRSWYL TENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDLSQLSVCLHEVAYWYILS
IGAQTDFLSVFFSGYTFKHKMVEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRG
MTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSFQSNVSNATNPPVLKRHRQ
EITRRTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWD
YGMSSSPHVLNRNRAQSGSVPQFKKVVFQFTDGSFTQPLYRGELNEHLGLLGPYIRAEVE
DNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPKKNFVKPNETKTYFWKVQHMAPTKD
EFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKS
WYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQIRWYLLSMG
SNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPKAGIWRVECLIGEHLHA
GMSTLFLVYSNKCQTPLGMASGHIRDFOITASGQYGQWAPKLARLHYSGSINAWSTKEPF
SWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFG
NVDSSGIKHNIFNPPIIARYIRLHPHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDA
QITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGV
KSLTSMYVKEFLISSQDGHQWTLFFQNGKVVFQGNQDSFTPVNSLDPPLLTRYLRI
HPQSWVHQIALRMEVLGCEAQDLY (SEQ ID NO:2)

Hình 2

4/28

CS12-CRM8.2-Vr

ttggccaactccctctctgcgcgctcgctcgctcactgaggccgggaccaaaggctcgcccgacgcc
 cgggctttgcccgggaggcctcagtgcgcgagcgagcgagagaggagtgcccaactccatcac
 taggggttcctggggaggctgctggtgaatattaaccaaggtcacccagttatcgaggagcaaa
 caggggctaagtccaccggggaggctgctggtgaatattaaccaaggtcacccagttatcgagg
 agcaaacaggggctaagtccacaaatgacctattaagaatatttcatagaacgaatgttccgatgct
 ctaatctctctagacaagggtcatatttgtatgggttacttattctctctttgttgactaagtcaat
 aatcagaatcagcagggttgcagtcagattggcagggataagcagcctagctcaggagaagtgcagta
 taaaagccccaggctgggagcagccatcacagaagtcactcattcttggcaggccaccatgcagat
 tgagctgagcactgcttcttctgtgcctgctgaggttctgcttctctgccaccaggagatactac
 ctgggggctgtggagcttcttgggactacatgcagctgcacctgggggagctgcctgtggatgcca
 ggttcccaccagagtgcccaaactccttccattcaacacctctgtggtctacaagaagacctctt
 tgtggagttcactgaccacctgttcaacattgccaaacccaggccaccctggatgggactcctggga
 cccaccattcaggctgaggtgtatgacactgtggtcgctcaccctcaagaacatggcctcccacctg
 tgagcctgcatgctgtgggggtcagctactggaagtcctctgagggggctgagtatgatgaccagac
 ctcccagagggagaaggaggatgacaaagtgttccctgggaagagccacacctatgtgtggcaggtc
 ctcaaggagaatggccccactgcctctgacccacctgcctgacctactcctacctttctcatgtgg
 acctggtcaaggacctcaactctggactgattggggccctgctggtgtgcaggggagggtccctggc
 caaagagaagaccagacctgcacaagttcattctcctgtttgctgtctttgatgagggcaagagc
 tggcactctgaaaccaagaactccctgatgcaggacagggatgctgcctctgccaggggcctggcca
 agatgcacactgtgaatggctatgtgaacaggagcctgcctggactcattggctgccacaggaaatc
 tgtctactggcatgtgattggcatggggacaacccctgaggtgcactccattttcttggaggggcac
 accttctgtgtcaggaaccacagacaggccagcctggagatcagccccatcaccttctcactgcc
 agacctgctgatggacctcggacagttcctgctgttctgccacatcagctcccaccagcatgatgg
 catggaggcctatgtcaagggtggacagctgccctgaggagccacagctcaggatgaagaacaatgag
 gaggctgaggactatgatgatgacctgactgactctgagatggatgtggtccgctttgatgatgaca
 acagcccatccttcattcagatcagggtctgtggccaagaacaccccaagacctgggtgcactacat
 tgctgctgaggaggaggactgggactatgccccactggtcctggccccctgatgacaggagctacaag
 agccagtacctcaacaatggccccacagaggattggacgcaagtacaagaaagtcagggttcatggcct
 aactgatgaaaccttcaagaccaggggaggccattcagcatgagctctggcatcctgggccccactcct
 gtatggggagggtgggggacacctgctcatcatcttcaagaaccaggcctccaggccctacaacatc
 taccacatggcatcactgatgtcaggccccctgtacagccgcaggctgccaaggggggtgaaacacc
 tcaaggacttccccattctgctggggagatcttcaagtacaagtggaactgtcactgtggaggatgg
 accaaccaaatctgacccaggtgcctcaccagatactactccagctttgtgaacatggagagggac
 ctggcctctggcctgattggccactgctcatctgtctacaaggagtctgtggaccagaggggaaacc
 agatcatgtctgacaagaggaatgtgattctgttctctgtctttgatgagaacaggagctggtacct
 gactgagaacattcagcgttctcctgccaacccctgctggggtgcagctggaggaccctgagttccag
 gccagcaacatcatgcaactccatcaatggctatgtgtttgacagcctccagctttctgtctgcctgc
 atgaggtggcctactggtacattctttctattggggcccagactgacttcttctgtcttcttctc
 (Tiếp)

Hình 3A

5/28

tggctacaccttcaaacacaagatgggtgtatgaggacaccctgacccctcttcccattctctggggag
 actgtgttcatgagcatggagaaccctggcctgtggattctgggatgccacaactctgacttccgca
 acaggggcatgactgccctgctcaaagtctcctcctgtgacaagaacactggggactactatgagga
 cagctatgaggacatctctgcctacctgctcagcaagaacaatgccattgagcccaggagcttcagc
 cagaatgtgagcaataatgccaccaac
 ccacctgtcctgaaacgccaccagaggagatcaccaggaccaccctccagtctgaccaggaggaga
 ttgactatgatgacaccatttctgtggagatgaagaaaggagactttgacatctatgacgaggacga
 gaaccagagcccaaggagcttccagaagaagaccaggcactacttcattgctgctgtggagcgctg
 tgggactatggcatgagctccagcccccatgtcctcaggaacagggcccagtctggctctgtgccac
 agttcaagaaagtggcttccaagagttcactgatggcagcttcacccagccccgttacagagggga
 gctgaatgagcacctgggactcctgggcccatacatcagggctgaggtggaggacaacatcatggtg
 accttccgcaaccaggcctccaggccctacagcttctacagctccctcatcagctatgaggaggacc
 agaggcagggggctgagccacgcaagaactttgtgaaacccaatgaaaccaagacctacttctggaa
 agtccagcaccacatggccccaccaaggatgagtttgactgcaaggcctgggcctacttctctgat
 gtggacctggagaaggatgtgcactctggcctgattggcccactcctggtctgccacaccaacccc
 tgaacctgcccattggaaggcaagtgactgtgcaggagtttgccctcttcttcaccatctttgatga
 aaccaagagctggtacttctactgagaacatggagcgcaactgcagggccccatgcaacattcagatg
 gaggacccccaccttcaaagagaactaccgcttccatgccatcaatggctacatcatggacacctgc
 ctgggcttgtcatggcccaggaccagaggatcaggtggtacctgcttctatgggctccaatgagaa
 cattcactccatccacttctctgggcatgtcttcaactgtgcgcaagaaggaggagtacaagatggcc
 ctgtacaacctctacctggggtctttgagactgtggagatgctgccctccaaagctggcatctgga
 ggggtggagtgcctcattgggggagcacctgcatgctggcatgagcacctgttctggtctacagcaa
 caagtgccagacccccctgggaatggcctctggccacatcagggaacttcagatcactgcctctggc
 cagtatggccagtggggcccccaagctggccaggtccactactctggatccatcaatgcctggagca
 ccaaggagccattcagctggatcaaagtggacctgctggcccccatgatcatccatggcatcaagac
 ccagggggccaggcagaagttctccagcctgtacatcagccagttcatcatcatgtacagcctggat
 ggcaagaaatggcagacctacagaggcaactccactggaacactcatggtcttctttggcaatgtgg
 acagctctggcatcaagcacaacatcttcaaccccccaatcatcgccagatacatcaggctgcaccc
 caccactacagcatccgcagcacccctcaggatggagctgatgggctgtgacctgaactcctgcagc
 atgcccctgggcatggagagcaaggccatttctgatgccagatcactgcctccagctacttcacca
 acatgtttggccacctggagcccaagcaaggccaggctgcacctccagggaaggagcaatgcctggag
 gccccaggtcaacaacccaaaggagtggctgcagggtggacttccagaagaccatgaagggtcactggg
 gtgaccacccaggggggtcaagagcctgctcaccagcatgtatgtgaaggagtctcctgatcagctcca
 gccaggatggccaccagtggacctcttcttccagaatggcaagggtcaagggtgttccagggcaacca
 ggacagcttccccctgtggtgaacagcctggacccccccctcctgaccagatacctgaggattcac
 ccccagagctgggtccaccagattgccctgaggatggaggtcctgggatgtgaggcccaggacctgt
 actgatga
 aataaaaagatctttattttcattagatctgtgtgttgggtttttgtgtgaggaacccctagtgtatgg
 Agttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaagggtcgcccgacg
 cccgggctttgcccggggcctcagtgcgagcgagcgcgagagaggagtggtggccaa (SEQ
 ID NO: 3)

Hình 3B

6/28

5'-ITR

ttggccaactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggctcgcccgacgcc
cgggctttgcccgggcggcctcagtgagcgagcgagcgcgagagagggagtgccaactccatcac
taggggttcct (SEQ ID NO:4)

CRM8

gggggaggctgctggtgaatattaaccaagggtcaccccagttatcgaggagcaaacaggggctaag
tccac (SEQ ID NO:5)

hTTR-Pro

aatgacctattaagaatatttcatagaacgaatgttccgatgctctaattctctctagacaagggtc
atatttgatgggttacttattctctctttgttgactaagtcaataatcagaatcagcaggtttgca
gtcagattggcagggataagcagcctagctcaggagaagtgagtataaaagccccaggctgggagca
gccatcacagaagtccactcattcttggcagg (SEQ ID NO:6)

Kozak

ccacc (SEQ ID NO:7)

Poly-A

aataaaagatctttattttcattagatctgtgtgttggtttttgtgtg (SEQ ID NO:8)

3'-ITR

aggaaccctagtgatggAgttggccaactccctctctgcgcgctcgctcgctcactgaggccgggcg
accaaaggctcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcgagagag
ggagtggccaa (SEQ ID NO:9)

Hình 4

7/28

CS12-CRM8.2-Vrp

ttggccaactccctctctgcgcgctcgctcgctcactgagggcgggcgaccaaaggtcgcccgacgco
 cgggctttgcccgggcgccctcagtgcgagcgagcgcgagagaggagtggccaactccatcac
 taggggttcctgggggaggtgctggtgaatattaaccaaggtcaccctcagttatcgaggagcaaa
 caggggctaagtcaccgggggaggtgctggtgaatattaaccaaggtcaccctcagttatcgagg
 agcaaacaggggctaagtcacaaatgacctattaagaatatttcatagaacgaatgttccgatgct
 ctaatctctctagacaagggttcataatttgatgggttacttattctctctttgttgactaagtcaat
 aatcagaatcagcaggtttgcagtcagattggcagggataagcagcctagctcaggagaagtgcagta
 taaaagccccaggtgaggagcagccatcacagaagtcactcattcttgaggccaccatgcagat
 tgagctgagcacctgcttcttctgctgctgaggttctgcttcttgccaccaggagatactac
 ctgggggctgtggagcttcttgggactacatgcagctgcagctgggggagctgcctgtggatgcca
 ggttcccaccagagtgcccaaatccttccattcaaacctctgtggtctacaagaagacctctt
 tgtggagttcactgaccacctgttcaacattgccaaaccagggccacctggatgggactcctggga
 cccaccattcaggtgaggtgtatgacactgtggtcgtcaccctcaagaacatggcctcccacctg
 tgagcctgcatgctgtgggggtcagctactggaagtcctctgaggggctgagatgatgaccagac
 ctcccagagggagaaggaggtgacaaagtgttccctgggaagagccacacctatgtgtggcaggtc
 ctcaaggagaatggccccactgcctctgacccacctgcctgacctactcctaccttctcatgtgg
 acctggtcaaggacctcaactctggactgattggggccctgctggtgtgcagggagggctccctggc
 caaagagaagaccagacctgcacaagttcattctcctgtttgctgtctttgatgagggcaagagc
 tggcactctgaaaccaagaactccctgatgcaggacagggatgctgcctctgcccagggcctggcca
 agatgcacactgtgaatggctatgtgaacaggagcctgcctggactcattggctgccacaggaaatc
 tgtctactggcatgtgattggcatggggacaacacctgaggtgcactccatttctctggagggccac
 accttctggtcaggaaccacagacagggccagcctggagatcagccccatcaccttctcactgccc
 agacctgctgatggacctcggacagttcctgctgttctgccacatcagctcccaccagcatgatgg
 catggaggcctatgtcaaggtggacagctgcctgaggagccacagctcaggatgaagaacaatgag
 gaggtgaggactatgatgatgacctgactgactctgagatggatgtggtccgctttgatgatgaca
 acagcccatccttcattcagatcaggtctgtggccaagaaacaccccaagacctgggtgcaactacat
 tgctgctgaggaggaggactgggactatgccccactggctcctggccccctgatgacaggagctacaag
 agccagtacctcaacaatggcccacagaggattggagcgaagtacaagaaagtcaggttcatggcct
 aactgatgaaaccttcaagaccaggaggccattcagcatgagctctggcatcctgggcccactcct
 gtatggggaggtgggggacacctgctcatcatcttcaagaaccaggcctccaggccctacaacatc
 taccacatggcatcactgatgtcaggccctgtacagccgcaggttgccaaagggggtgaaacacc
 tcaaggacttccccattctgcctggggagatcttcaagtacaagtggactgtcactgtggaggatgg
 accaaccaaatctgacccaggtgcctcaccagatactactccagctttgtgaacatggagagggac
 ctggcctctggcctgattggcccactgctcatctgctacaaggagtctgtggaccagaggggaaacc
 agatcatgtctgacaagaggaatgtgattctgttctctgtctttgatgagaacaggagctggtacct
 gactgagaacattcagcgcttctgcccacacctgctggggtgcagctggaggacctgagttccag
 gccagcaacatcatgcactccatcaatggctatgtgtttgacagcctccagctttctgtctgcctgc
 atgaggtggcctactggtacattctttctattggggcccagactgacttctttctgtcttcttctc
 (Tiếp)

Hình 5A

8/28

tggctacaccttcaaacacaagatgggtgtatgaggacacccctgacccctcttcccattctctggggag
 actgtgttcatgagcatggagaaccctggcctgtggattctgggatgccacaactctgacttccgca
 acaggggcatgactgccctgctcaaagtctcctcctgtgacaagaacactggggactactatgagga
 cagctatgaggacatctctgcctacctgctcagcaagaacaatgccattgagcccaggagcttcagc
 cagaatgtgagcaataatgccaccaaccacctgtcctgaaacgccaccagaggagatcaccagga
 ccaccctccagtctgaccaggaggagattgactatgatgacaccatttctgtggagatgaagaaaga
 ggactttgacatctatgacgaggagcagaaccagagcccaaggagcttccagaagaagaccaggcac
 tacttcattgctgctgtggagcgcctgtgggactatggcatgagctccagccccatgtcctcagga
 acagggcccagctctggctctgtgccacagttcaagaaagtggctcttccaagagttcactgatggcag
 cttcaccacagccccctgtacagaggggagctgaatgagcacctgggactcctgggccatacatcagg
 gctgaggtggaggacaacatcatggtgaccttccgcaaccaggccctccaggccctacagcttctaca
 gctccctcatcagctatgaggaggaccagaggcagggggctgagccacgcaagaactttgtgaaacc
 caatgaaaccaagacctacttctggaaagtccagcaccacatggccccaccaaggatgagtttgac
 tgcaaggcctgggctacttctctgatgtggacctggagaaggatgtgcactctggcctgattggcc
 cactcctggtctgccacaccaacacctgaacctgcccattggaaggcaagtgactgtgcaggagtt
 tgccctcttcttcaccatctttgatgaaaccaagagctggtacttactgagaacatggagcgcacac
 tgcagggccccatgcaacattcagatggaggacccccaccttcaaagagaactaccgcttccatgcca
 tcaatggctacatcatggacacctgacctgggcttgcattggccaggaccagaggatcaggtggta
 cctgctttctatgggctccaatgagaacattcactccatccacttctctgggcatgtcttactgtg
 cgcaagaaggaggagtacaagatggccctgtacaacctctacctggggcttttgagactgtggaga
 tgctgccctccaaagctggcatctggaggggtggagtgcctcattgggggagcacctgcatgctggcat
 gagcacctgttccctgggtctacagcaacaagtgccagacccccctgggaatggcctctggccacatc
 agggacttccagatcactgcctctggccagtatggccagtgggcccccaagctggccaggctccact
 actctggatccatcaatgcctggagcaccaaggagccattcagctggatcaaagtggacctgctggc
 cccatgatcatccatggcatcaagacccagggggccaggcagaagttctccagcctgtacatcagc
 cagttcatcatcatgtacagcctggatggcaagaaatggcagacctacagaggcaactccactggaa
 cactcatggctcttctttggcaatgtggacagctctggcatcaagcacaacatcttcaaccccccaat
 catcgccagatacatcaggctgcacccccccactacagcatccgcagcacctcaggatggagctg
 atgggctgtgacctgaactcctgcagcatgccccctgggcatggagagcaaggccatttctgatgcc
 agatcactgcctccagctacttcaccaacatgtttgccacctggagcccaagcaaggccaggctgca
 cctccagggaaggagcaatgcctggaggccccaggtcaacaacccaaaggagtggtgaggtggac
 ttccagaagaccatgaaggctcactgggggtgaccaccagggggtcaagagcctgctcaccagcatgt
 atgtgaaggagttcctgatcagctccagccaggatggccaccagtggaacctcttcttccagaatgg
 caaggtaagggtgttccagggaaccaggacagcttccccctgtggtgaacagcctggaccccccc
 ctctgaccagatacctgaggattcacccccagagctgggtccaccagattgcctgaggatggagg
 tcttgggatgtgaggcccaggacctgtactgatgaaataaaagatctttattttcattagatctgtg
 tgttgggtttttgtgtgaggaacccctagtgtggAgttggccactccctctctgcgcgctcgctcg
 ctactgaggccggggcgaccaaaggctcgcccgacgcccgggctttgcccggggcgccctcagtgagcg
 agcgagcgcgagagaggagtggtgccaagacgatttaaatgacaagcttggcgtaatcatggtcata
 gctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacagagccggaagcataaag
 tgtaaagCctgggggtgcctaataatgagttagtaactcacattaattgcgttgcgctcactgcccgtt
 tccagtcgggaaacctgtcggtgccagctgcattaatgaatcggcccaacgcgcggggagaggcggtt

(Tiếp)

Hình 5B

gcgatttgggcgtcttccgcttccgctcgtcactgactgcgtgcgttcggtcgttcggttcggtcgcgcga
gcggtatcagctcactcaaaaggcggtaatcacggttatccacagaatcaggggataacgcaggaaaga
acatgtgagcaaaaaggccagcaaaaaggccaggaaccgtaaaaaggcgcgttgcgtggcggtttttcca
taggtccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaaaccgcaca
ggactataaagataaccaggcggtttccccctggaagctccctcgtgcgctctcctgttccgacctgc
cgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctg
taggtatctcagttcgggtgtaggtcggttcgctccaagctgggctgtgtgcacgaaccccccggttcag
cccgaccgctgcgccttatccggtaactatcgtcttgagtcacaaccccggttaagacacgacttatcgc
cactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttctt
gaagtgggtggcctaactacggctacactagaagaacagtatatttggtatctgcgctctgctgaagcca
gttaccttcggaaaaagagttggtagctcttgatccggcaaaacaaaccacgcgtggtagcgggtgggt
tttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttc
tacgggggtctgacgctcagtggaacgaaaaactcacgttaagggatttttggctatgagattatcaaaa
aggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagt
aaacttgggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcg
ttcatccatagttgcctgactccccgctcgtgtagataactacgatacgggaggggttaccatctggc
cccagtgctgcaatgataaccgcgagacccacgctcacccggtccagatttatcagcaataaaccagc
cagccggaagggcgagcgcagaagtggctcctgcaactttatccgcctccatccagtcctattaattg
ttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacggtgttgccattgctaca
ggcatcgtgggtgtcacgctcgtcggttggtatggcttcattcagctccgggttcccaacgatcaaggc
gagttacatgatcccccatggttggtgcaaaaaagcgggttagctccttcggctcctccgatcgttgtcag
aagtaagttggcgcagtggttatcactcatggttatggcagcactgcataattctcttactgtcatg
ccatccgtaagatgctttttctgtgactggtgagtagtactcaaccaagtcattctgagaatagtgtatgc
ggcgaccgagttgctcttgcccggtcgaatacgggataataccgcgccacatagcagaactttaaa
agtgtcatcattggaaaacggttcttcggggcgaaaactctcaaggatcttaccgctgttgagatcc
agttogatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccageggtttctg
ggtgagcaaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatggtgaat
actcactcttccctttttcaatattattgaagcatttatcagggttattgtctcatgagcgggatac
atatttgaatgtatttagaaaaataaacaataagggttccgcgcacattttccccgaaaagtgccac
ctgacgtctaagaaaccattattatcatgacattaacctataaaaaataggcggtatcacgagggcctt
tcgtctcgcgcgttttcggtgatgacgggtgaaaacctctgacacatgcagctcccggagacgggtcaca
gcttgctgtgaagcggatgccgggagcagacaagcccgtcagggcgcgctcagcgggtgttggcggggt
gtcgggggtggttaactatgcccgcacagagcagattgtactgagagtgaccatattgcgggtgtga
aataccgcacagatgcgtgaaggagaaaaataccgcacatcaggcgccatttcgccattcaggetgcgcac
tgttgggaagggcgatcgggtgcgggctcttcgctattacgccagctggcgaaaggggggatgtgctg
caaggcgattaagttgggtaacgccagggttttccagtcacgacggttgtaaaacgacggccagtg
attcctcgagatttaaatgacg (SEQ ID NO:10)

10/28

CS12-CRM8.2-V

ttggccactccctctctgcgcgctcgctcgctcactgagggcgggacgacaaaggctgcccgcagcc
 cgggctttgcccgggcgccctcagtgcgagcgagcgagcgagagggagtgcccaactccatcac
 taggggttcctgagtttaaaacttcgtcgacgggggaggctgctggtgaatattaaccaaggtcaccc
 cagttatcggaggagcaaacaggggctaagtccacgggggaggctgctggtgaatattaaccaagg
 tcaccccagttatcggaggagcaaacaggggctaagtccacaaatgacctattaagaatatttcata
 gaacgaatgttccgatgctcctaattctctctagacaaggttcataatttgatgggttacttattctct
 ctttggtgactaagtcaataatcagaatcagcagggttgacagtcagattggcagggtataagcagcct
 agctcaggagaagtgcgtataaaagccccaggctgggagcagccatcacagaagtccactcattctt
 ggcaggggcgcgccgccaccatgcagattgagctgagcacctgcttcttctgctgctgaggtt
 ctgcttctctgccaccaggagatactacctgggggctgctggagccttcttgggactacatgcagctt
 gacctgggggagctgctgctggtatgccagggttcccacccagagtgcccaaactccttccattcaaca
 cctctgtggtctacaagaagaccctcttctgtggagttcactgaccacctgttcaacattgccaacc
 caggccaccctggatgggactcctgggacccaccattcaggctgaggtgtatgacactgtggctgct
 accctcaagaacatggcctcccaccctgtgagcctgcattgctgtgggggtcagctactggaagtcct
 ctgagggggctgagtatgatgaccagacctcccagagggagaaggaggatgacaaagtgttccctgg
 gaagagccacacctatgtgtggcaggctcctcaaggagaaatggccccactgcctctgacccaccctgc
 ctgacctactcctaccttctctcatgttggaacctggtcaaggacctcaactctggactgattggggccc
 tgcctggtgtgcagggagggtcctcctggccaaagagaagacccagacctgcacaagttcattctcct
 gtttgcctgtctttgatgagggcaagagctggcactctgaaaccaagaactccttgatgcaggacagg
 gatgctgcctctgccagggcctggcccaagatgcacactgtgaatggctatgtgaacaggagcctgc
 ctggactcatttgctgccacaggaaatctgtctactggcatgtgattggcatggggacaacccctga
 ggtgcactccattttcctggaggggccacaccttctggtcaggaaccacagacaggccagcctggag
 atcagccccatcaccttccctcactgcccagacctgctgatggacctcgacagttcctgctgttct
 gccacatcagctcccaccagcatgatggcatggaggcctatgtcaagggtggacagctgcctgagga
 gccacagctcaggatgaagaacaatgaggaggctgaggactatgatgatgacctgactgactctgag
 atggatgtgtgctccgctttgatgatgacaacagccccatccttcattcagatcaggtctgtggccaaga
 aacacccccaaagacctgggtgcactacattgctgctgaggaggaggactgggactatgccccactggt
 cctggccccctgatgacaggagctacaagagccagctacctaacaatggccccacagaggattggacgc
 aagtacaagaagtcagggttcatggcctacactgatgaaaccttcaagaccagggaggccattcagc
 atgagctctggcatcctggggccactcctgtaaggggagggtgggggacacctgctcatcatcttcaa
 gaaccaggcctccaggccctacaacatctaccacatggcatcactgatgtcaggccctgtacagc
 cgcaggctgccaaaggggggtgaaacacctcaaggacttccccattctgcctggggagatcttcaagt
 acaagtggactgtcactgtggaggatggaccaaccaaactctgaccccagggtgcctcaccagatacta
 ctccagctttgtgaacatggagagggacctggcctctggcctgattggcccactgctcatctgctac
 aaggagtctgtggaccagaggggaaaccagatcatgtctgacaagagggaatgtgattctgttctctg
 tctttgatgagaacaggagctggtacctgactgagaacattcagcgttctctgcccacccctgctgg
 ggtgcagctggaggaccctgagttccaggccagcaacatcatgcactccatcaatggctatgtgttt
 gacagcctccagctttctgtctgctgcatgaggtggcctactggtacattcttcttattggggccc
 agactgacttcccttctgtcttcttctctggctacaccttcaaacacaagatgggtgtatgaggacac
 cctgacctcttcccatctctgtggggagactgtgttcatgagcatggagaaccctggcctgtggatt
 ctgggatgccacaactctgacttccgaacaggggcatgactgcctgctcaaagtcctcctctgtg
 (Tiếp)

Hình 6A

11/28

acaagaacactggggactactatgaggacagctatgaggacatctctgcctacctgctcagcaagaa
 caatgccattgagcccaggagcttcagccagaatgtgagcaataatgccaccaacccacctgtcctg
 aaacgccaccagagggagatcaccaggaccacccctccagtctgaccaggaggagattgactatgatg
 acaccatcttctgtggagatgaagaaagaggactttgacatctatgacgaggacgagaaccagagccc
 aaggagcttccagaagaagaccaggcactacttccattgctgctgtggagcgcctgtgggactatggc
 atgagctccagccccatgtcctcaggaacaggggccagctctggctctgtgccacagttcaagaaag
 tggctctccaagagttcactgatggcagcttcacccagccccgtacagaggggagctgaatgagca
 cctgggactcctgggcccatacatcagggtgaggtggaggacaacatcatggtgaccttccgcaac
 caggcctccaggccctacagcttctacagctccctcatcagctatgaggaggaccagaggcaggggg
 ctgagccacgcaagaactttgtgaaacccaatgaaaccaagaccctacttctggaaagtccagacca
 catggcccccaccaaggatgagtttgactgcaaggcctgggcctacttctctgatgtggacctggag
 aaggatgtgcactctggcctgattggcccactcctggctctgccacaccaacacccctgaacctgccc
 atggaaggcaagtgactgtgcaggagtttgccctcttcttcacatctttgatgaaaccaagagctg
 gtacttactgagaacatggagcgcgaactgcagggcccccattgcaacattcagatggaggacccacc
 ttcaaagagaactaccgcttccatgccatcaatggctacatcatggacacccctgcctgggcttgtca
 tggcccaggaccagaggatcagggtggtacctgcttctatgggctccaatgagaacattcactccat
 ccacttctctgggcatgtcttctactgtgcgcaagaaggaggagtacaagatggccctgtacaacctc
 taccctggggctctttgagactgtggagatgtctgccctccaaagctggcatctggaggggtggagtgcc
 tcatgggggagcacctgcatgtctggcatgagcaccctgttccctggctctacagcaacaagtgccagac
 cccctggggaatggcctctggccacatcagggaacttccagatcactgcctctggccagtatggccag
 tgggcccccaagctggccaggctccactactctggatccatcaatgcctggagcaccaaggagccat
 tcagctggatcaaagtggacctgctggcccccatgatcatccatggcatcaagacccagggggccag
 gcagaagttctccagcctgtacatcagccagttcatcatcatgtacagcctggatggcaagaaatgg
 cagacctacagaggcaactccactggaacactcatggtcttctttggcaatgtggacagctctggca
 tcaagcacaacatcttcaaccccccaatcatcgccagatacatcaggctgcacccacccactacag
 catccgcagcacccctcaggatggagctgatgggctgtgacctgaactcctgcagcatgcccctgggc
 atggagagcaaggccatcttctgatgccagatcactgcctccagctacttcaccaacatgtttgcca
 cctggagcccaagcaaggccaggctgcacctccagggaaggagcaatgcctggaggccccagggtcaa
 caacccaaaggagtggctgcagggtggacttccagaagaccatgaagggtcactgggggtgaccacccag
 ggggtcaagagcctgtctcaccagcatgtatgtgaaggagtccctgatcagctccagccaggatggcc
 accagtggaccctcttcttccagaatggcaagggtcaagggtgttccagggcaaccaggacagcttcac
 ccctgtggtgaacagcctggacccccccctcctgaccagatacctgaggattcacccccagagctgg
 gtccaccagattgcctgaggatggagggtcctgggatgtgaggccaggacctgtactgatgagcgg
 ccgctcttagtagcagtatcgataataaaagatctttatcttatttagatctgtgtgtgtgtgtgtgt
 gtgtgttaattaagctcgcaagggaacccctagtgtggAgttggccactccctctctgcgcgctcg
 ctcgctcactgaggccgggacgacaaagggtcgcccgacgcccgggctttgcccgggcccgcctcagt
 agcgagcgcgcgcgcagagagggagtggccaa (SEQ ID NO:38)

Hình 6B

12/28

CS04-FL-NA

atgcagattgagctgagcacctgcttcttctgtgcctgctgaggttctgcttctctgccaccagga
gatactacctgggggctgtggagcttcttgggactacatgcagcttgacctgggggagctgacctgt
ggatgccaggttcccacccagagtgcccaaattccttcccatccaacacctctgtggtctacaagaag
acctctttgtggagttcactgaccacctgttcaacattgccaaaccaggccacctggatgggac
tctggggacccaccattcaggctgaggtgtatgacactgtgggtcatcacctcaagaacatggcctc
ccacctgtgagcctgcatgctgtgggggtcagctactggaaggcctctgagggggctgagtatgat
gaccagacctcccagagggagaaggaggatgacaaagtgttccctgggggcagccacacctatgtgt
ggcaggtcctcaaggagaatggcccatggcctctgacccactctgcctgacctactcctacctttc
tcatgtggacctggtcaaggacctcaactctggactgattggggccctgctggtgtgcagggagggc
tccctggccaaagagaagacccagacctgcacaagttcattctcctgtttgctgtctttgatgagg
gcaagagctggcactctgaaaccaagaactccctgatgcaggacagggatgctgacctctgccagggc
ctggcccaaagatgcacactgtgaatggctatgtgaacaggagcctgcctggactcattggctgccac
aggaaatctgtctactggcatgtgattggcatggggacaacccctgaggtgacctcattttctctgg
agggccacaccttctggtcaggaaccacagacagggccagcctggagatcagcccatcaccttct
cactgccagacctgctgatggacctcggacagttcctgctgttctgccacatcagctcccaccag
catgatggcatggaggcctatgtcaaggtggacagctgcctgaggagccacagctcaggatgaaga
acaatgaggaggctgaggactatgatgatgacctgactgactctgagatggatgtgggtccgctttga
tgatgacaacagcccatccttcattcagatcaggtctgtggccaagaaacaccccaagacctgggtg
cactacattgctgctgaggaggaggactgggactatgcccactggtcctggccctgatgcaggga
gctacaagagccagtacctcaacaatggccacagaggattggacgcaagtacaagaaagtccaggtt
catggcctacactgatgaaaccttcaagaccagggaggccattcagcatgagctgtggcatcctgggc
ccactcctgtatggggaggtgggggacacctgctcatcatcttcaagaaccaggcctccaggccct
acaacatctaccacatggcatcactgatgtcaggccctgtacagccgaggtgcacaaagggggt
gaaacacctcaaggacttccccattctgcctggggagatcttcaagtacaagtggaactgtcactgtg
gaggatggaccaaccaaatctgacccaggtgcctcaccagataactactccagctttgtgaacatgg
agaggacctggcctctggcctgattggcccatgctcatctgctacaaggagtctgtggaccagag
gggaaaccagatcatgtctgacaagaggaatgtgattctgttctctgtctttgatgagaacaggagc
tggtacctgactgagaacattcagcgttctcctgccaaacctgctggggtgcagctggaggacctg
agttccaggccagcaacatcatgcactccatcaatggctatgtgtttgacagcctccagctttctgt
ctgcctgcatgagggtggcctactgggtacattctttctattggggcccagactgacttctttctgtc
ttcttctctggctacaccttcaaacacaagatgggtgtatgaggacacctgacctcttcccattct
ctggggagactgtgttcatgagcatggagaacctggcctgtggattctgggatgccacaactctga
cttccgcaacaggggcatgactgccctgctcaaagtctcctcctgtgacaagaacactggggactac
tatgaggacagctatgaggacatctctgcctacctgctcagcaagaacaatgccattgagcccagga
gcttcagccagaatccacctgtcctgaaacgccaccagagggagatcaccaggaccacctccagtc
tgaccaggaggagattgactatgatgacaccatttctgtggagatgaagaaagaggactttgacatc
tatgacgaggacgagaaccagagcccaaggagcttcagaagaagaccaggcactacttcatgtgtg
ctgtggagcgctgtgggactatggcatgagctccagccccatgtcctcaggaacagggcccagtc
tggtctctgtgccacagttcaagaaagtgttcttccaagagttcactgatggcagcttcaccagccc
ctgtacagaggggagctgaatgagcacctgggactcctggggccatacatcagggtgaggtggagg
acaacatcatggtgaccttccgcaaccaggcctccaggccctacagcttctacagctccctcatcag
ctatgagGaggaccagagggcagggggctgagccacgcaagaactttgtgaaacccaatgaaaccaag
acctacttctggaaagtccagcaccacatggcccccaccaaggatgagtttgactgcaaggcctggg

(Tiếp)

Hình 7A

13/28

cctacttctctgatgtggacctggagaaggatgtgcactctggcctgattggcccactcctggtc
tgccacaccaacacctgaacctgcccattggaaggcaagtgactgtgcaggagtttgccctctt
cttcaccatctttgatgaaaccaagagctgggtacttcactgagaacatggagcgcaactgcaggg
cccatgcaacattcagatggaggacccaccttcaaagagaactaccgcttccatgccatcaat
ggctacatcatggacacctgacctgggcttgctcatggcccaggaccagaggatcaggtggtacct
gctttctatgggctccaatgagaacattcactccacttctctgggcatgtcttcaactgtgc
gcaagaaggaggagtacaagatggccctgtacaacctctacctggggctcttgagactgtggag
atgtgcccctccaaagctggcatctggagggtggagtgcctcattggggagcacctgcattggtg
catgagcacctgttccctgggtctacagcaacaagtgccagacccccctgggaatggcctctggcc
acatcagggacttccagatcactgcctctggccagtatggccagtgggcccccaagctggccagg
ctccactactctggatccatcaatgcctggagcaccaaggagccattcagctggatcaaagtgga
cctgctggcccccatgatcatccatggcatcaagaccagggggccaggcagaagttctccagcc
tgtacatcagccagttcatcatcatgtacagcctggatggcaagaaatggcagacctacagaggc
aactccactggaacactcatggtcttctttggcaatgtggacagctctggcatcaagcacaacat
cttcaaccccccaatcatcgccagatacatcaggctgcacccccaccactacagcatccgcagca
ccctcaggatggagctgatgggctgtgacctgaactcctgcagcatgcccctgggcatggagagc
aaggccatttctgatgccagatcactgcctccagctacttcaccaacatgtttgccacctggag
cccaagcaaggccaggctgcacctccagggaaggagcaatgcctggaggccccaggtcaacaacc
caaaggagtggctgcaggtggacttccagaagaccatgaaggctcactggggtgaccacccagggg
gtcaagagcctgctcaccagcatgtatgtgaaggagtctcctgatcagctccagccaggatggcca
ccagtggacctcttcttccagaatggcaagggtcaagggtgttccagggaaccaggacagcttca
cCctgtggtgaacagcctggacccccctcctgaccagatacctgaggattcacccccagagc
tgggtccaccagattgcctgaggatggagggtcctgggatgtgaggcccaggacctgtactga (S
EQ ID NO:37)

Hình 7B

NG1: V S N N V S N N A T N N (SEQ ID NO:11)
GTG AGC AAC AAT GTG AGC AAC AAT GCC ACC AAT AAT GCT ACC AAC (SEQ ID NO:12)

NG4: V S N N A T N N V S N (SEQ ID NO:13)
GTG AGC AAC AAT GCC ACC AAC AAT GTG AGC AAC (SEQ ID NO:14)

NG5: V S N N A T N (SEQ ID NO:15)
GTG AGC AAT AAT GCC ACC AAC (SEQ ID NO:16)

NG6: V S N N (SEQ ID NO:17)

GTG AGC AAT AAT (SEQ ID NO:18)

NG9: R S L (SEQ ID NO:19)
AGG AGC CTG (SEQ ID NO:20)

NG10: A T N V S N N S A T S A D S A V S (SEQ ID NO:21)
GCC ACT AAT GTG TCT AAC AAC TCT GCT ACC TCT GCT GAC TCT GCT GTG AGC (SEQ ID NO:22)

NG16: A T N Y V N R S L (SEQ ID NO:23)
GCC ACC AAC TAT GTG AAC AGG AGC CTG (SEQ ID NO:24)

Hình 8A

15/28

NG17: A T N Y V N R S L S A T S A D S A V S Q N (SEQ ID NO:25)
GCC ACC AAC TAT GTG AAC AGG AGC CTG TCT GCC ACC TCT GCT GAC TCT GCT GTG AGC CAG AAT (SEQ ID NO:26)

NG18: V S N N V S N A V S A V S A (SEQ ID NO:27)
GTG AGC AAC AAT GTG AGC AAT GCT GTG TCT GCT GTG TCT GCT (SEQ ID NO:28)

NG19: I T V A S A T S N I T V A S A D (SEQ ID NO:29)
ATC ACT GTG GCC TCT GCC ACC TCT AAC ATC ACT GTG GCC TCT GCT GAC (SEQ ID NO:30)

NG20: I T V T N I T V T A (SEQ ID NO:31)
ATC ACT GTG ACC AAC ATC ACT GTG ACT GCC (SEQ ID NO:32)

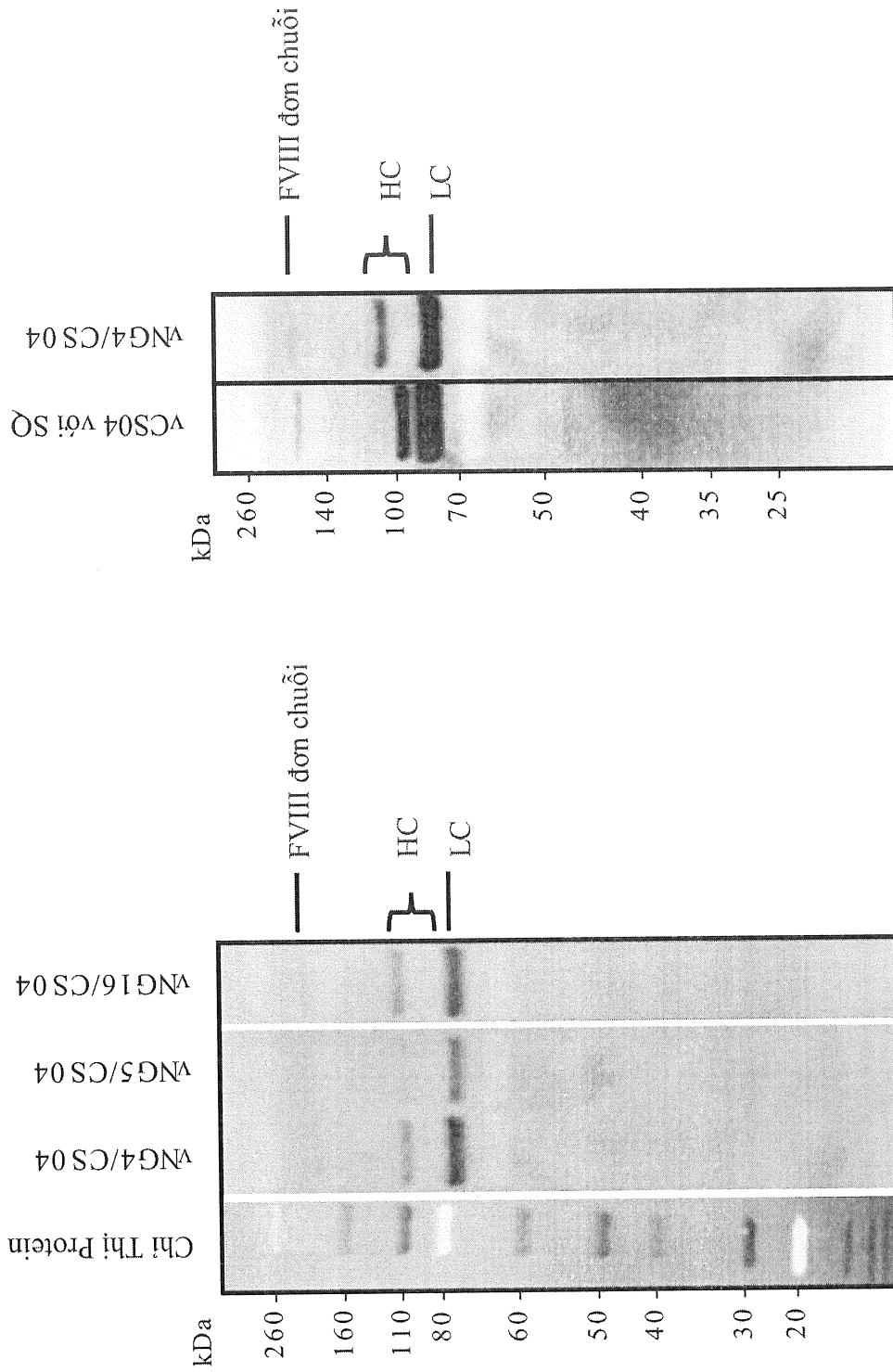
NG21: Q T V T N I T V T A (SEQ ID NO:33)

CAG ACT GTG ACC AAC ATC ACT GTG ACT GCC (SEQ ID NO:34)

NGV: A T N V S N N S N T S N D S N V S (SEQ ID NO:35)
GCC ACT AAT GTG TCT AAC AAC AGC AAC ACC AGC AAT GAC AGC AAT GTG TCT (SEQ ID NO:36)

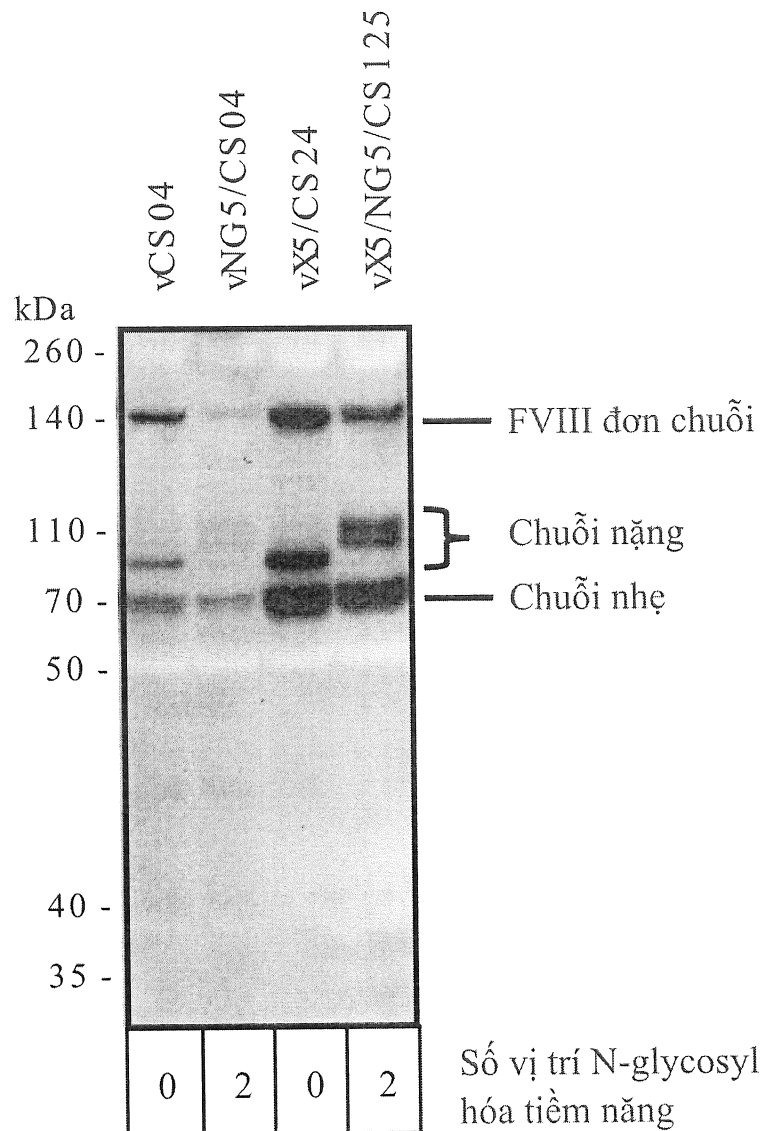
Hình 8B

16/28



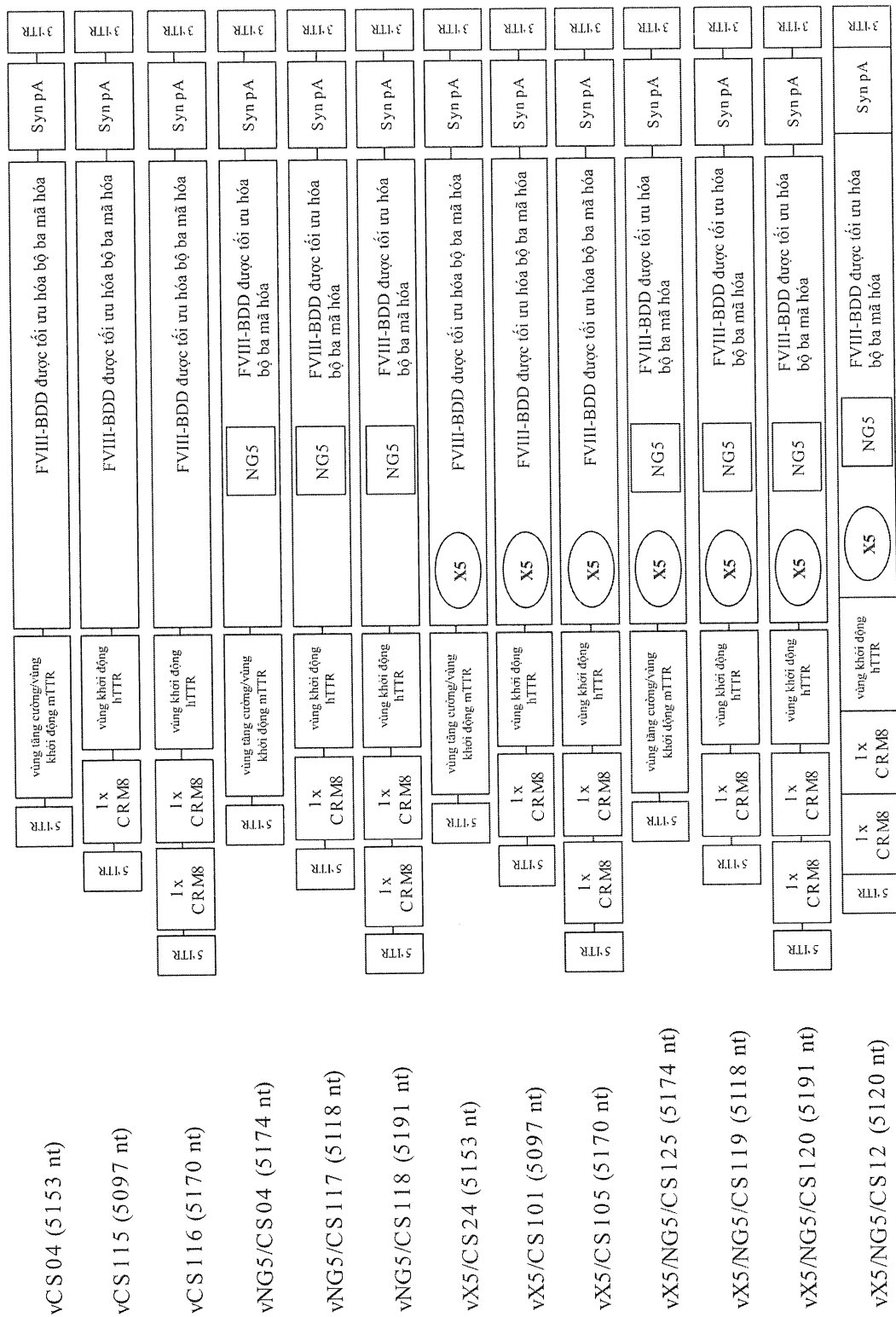
Hình 9

17/28



Hình 10

18/28

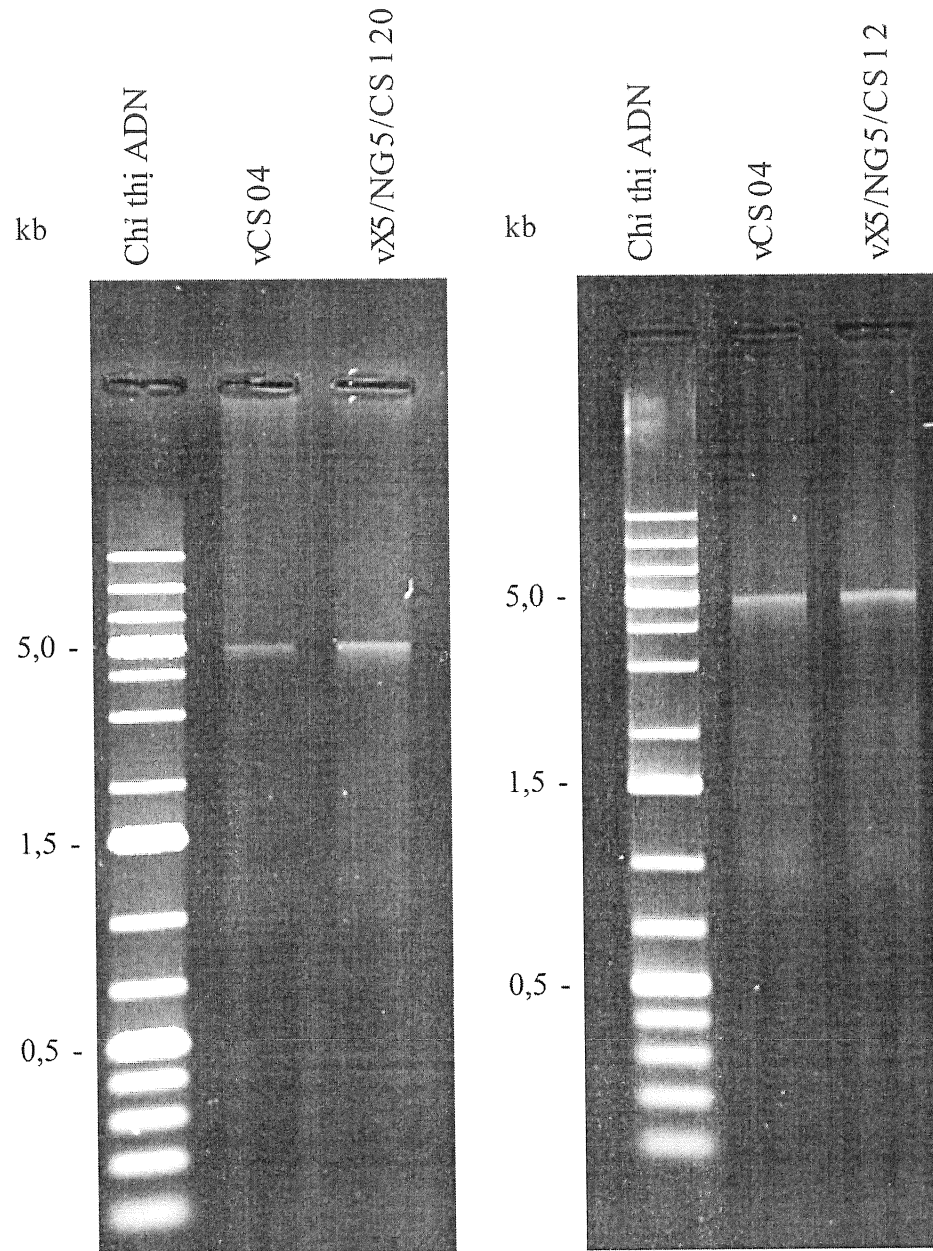


Hình 11

Vector	Vùng khởi động	CDS	(Các) cải biến trình tự	Hiệu lực sinh học in vivo (chuột nhắt “Đồng E”)				Hiệu lực sinh học in vitro (tế bào HepG2)		
				Số lượng chuột nhắt	Sự biểu hiện FVIII trung bình vào ngày 14 [IU/ml]	Độ lệch chuẩn	Số lần tăng lên của sự biểu hiện	BPU	Độ lệch chuẩn	Số lần tăng lên của sự biểu hiện*
vCS04	vùng tăng cường/vùng khởi động mTTR	CS04	-	44	2,1	1,7	1,0 (tham chiếu)	1,0	0,6	1,0 (tham chiếu)
vCS115	1xCRM8/hTTR	CS04	-	8	2,5	0,9	1,2	2,1	n.a.	2,2
vCS116	2xCRM8/hTTR	CS04	-	8	1,9	0,6	0,9	3,6	n.a.	3,7
vNG5/CS04	vùng tăng cường/vùng khởi động mTTR	CS04	NG5	8	4,0	2,6	1,9	0,6	n.a.	0,6
vNG5/CS117	1xCRM8/hTTR	CS04	NG5	7	3,2	1,0	1,5	1,5	0,1	1,5
vNG5/CS118	2xCRM8/hTTR	CS04	NG5	7	2,9	1,8	1,4	2,3	n.a.	2,4
vX5/CS24	vùng tăng cường/vùng khởi động mTTR	CS04	X5	16	6,5	2,1	3,1	3,8	1,9	4,0
vX5/CS101	1xCRM8/hTTR	CS04	X5	7	5,2	3,2	2,5	7,0	n.a.	7,3
vX5/CS105	2xCRM8/hTTR	CS04	X5	8	5,1	2,8	2,5	20,3	n.a.	21,2
vX5/NG5/CS125	vùng tăng cường/vùng khởi động mTTR	CS04	NG / X5	8	9,3	2,8	4,5	3,0	1,0	3,2
vX5/NG5/CS119	1xCRM8/hTTR	CS04	NG / X5	8	7,6	5,8	3,6	13,6	4,4	14,2
vX5/NG5/CS120	2xCRM8/hTTR	CS04	NG / X5	8	11,4	5,0	5,5	19,0	n.a.	19,8

Hình 12

20/28



Hình 13

Vector	Vùng khởi động	CDS	(Các) cải biến trình tự	Hiệu lực sinh học in vivo					Hiệu lực sinh học in vitro		
				Liều lượng [vg/kg]	Số lượng chuột nhất	Sự biểu hiện FV/III trung bình vào ngày 14 [IU/ml]	Độ lệch chuẩn	Số lần tăng lên của sự biểu hiện	BPU	Độ lệch chuẩn	Số lần tăng lên của sự biểu hiện*
vCS04	vùng tăng cường/vùng khởi động mTTR	CS04	-	1,0E+12	16	0,3	0,1	1,0*	1,0	0,1	1
				4,0E+12	16	2,5	0,9	-			
vX5/NG5/CS120	2xCRM8/hTTR	CS04	NG / X5	1,0E+12	7	4,3	1,7	14,5*	16,7	5,1	17
vX5/NG5/CS12	2xCRM8/hTTR	CS04	NG / X5, được làm giảm kích thước	5,0E+11	24	1,5	0,8	-	24,2	8,1	24
				1,0E+12	24	4,2	1,3	14,0*			

Hình 14

FVIII kiểu đại của người



BDD-FVIII kiểu Refacto



Hình 15

22/28

hFVIII-FL-AA

```

mqielstcfff lc1llrfcfssa trryylgave lswdymqsdl gelpvdarfp prvpksfpfn
tsvvykktlf veftdhlfni akprppwmgf lgptiqaevy dtvvitlknm ashpvs1hav
gvsywkaseg aeyddqtsqr ekeddkvfpg gshtyvwqvl kengpmasdp lcltysylsh
vdlvkdlng ligallvcre gslakektqt lkhfillfav fdegkswhse tkns1mqdrd
aasarawpkm htvngyvnrsl pgligchrk svywhvigmg ttpevhsifl eghtflvrnh
rqasleispi tfltaqtllm dlqqlfllch issqhhdgme ayvkvdscpe epqlrmknne
eaedydddltd dsemdvvrfd ddnspsfiqi rsvakkhpkt wvhyiaaeed dwdyaplvla
pddrsyksqy lnnqpqrigr kykkvrfmay tdetfktrea iqhesgilgp llygevgdtl
liifknqasr pyniypghit dvrplysrll pkgvkh1kdf pilpgeifky kwtvtvedgp
tksdprcltr yyssfvnmer dlasgligpl licykesvdq rgnqimsdkr nvilfsvfde
nrswylteni qrflpnpagv qledpefqas nimhsingyv fds1qlsvcl hevaywyils
igaqtdflsv ffsgytfkhk mvyedtltlf pfsgetvfms menpglwilg chnsdfrnrg
mtallkvssc dkntgdyied syedisayll sknnaieprs fsqnsrhpst rqkqfnatti
pendiektdp wfahrtmpk iqnvsssdll mllrqsptph glslsdlqea kyetfsddps
pgaidsnns1 semthfrpql hhsghdmvftp esglqlrlne klgttaatel kklfdkvsst
snnlistips dnlaagtdnt sslgppsmvp hydsqldt1 fgkksplte sggplslsee
nndskllesg lmnsgesswg knvsstesgr lfkghkrahgp alltkdnalf kvsisl1ktn
ktsnnsatnr kthidgpsll ienspsvwqn ilesdtefkk vtplihdrml mdknata1rl
nhmsnkt1ss knmemvqqk egpipdaqn pdmsffkmlf lpesarwiqr thgkns1nsg
qgspspkqlvs lgpeksvegq nflseknkvv vgkgeftkdv glkemvfpss rnlfltnldn
lhennthnqe kkiqeeiekk etliqenvvl pqihtvtgtk nfmkn1flls trqnvegsyd
gayapvlqdf rslndstnrt kkhtahfskk geeenleglg nqtkqiveky acttrispnt
sqqnfv1tqr kralqkfrlp leetelekri ivddtstqws knmkhltpst ltqidyneke
kgaitqsp1s dcltrshsip qanrsplpia kvssf1sirp iyltrvlfqd nsshlpaasy
rkkdsgvqes shflqgakn nls1lailtle mtgdqrevgs lgtsatnsvt ykkventvlp
kpd1pkts1k vellpkvhiy qkdlfptets ngspghldlv egsl1qgteg aikwneanrp
gkvpflrvat essaktpskl ldplawdnhy gtqipkeewk sqekspekta fkkkdt1sl
nacesnhaia ainegqnkpe ievtwakqgr terlcsqnpp vlkrhqreit rtt1qsdqee
idyddtisve mkkedfdiyd edenqsprsf qkktrhyfia averlw1dygm sssphvlrnr
aqsgsvpqfk kvvfqeftdg sftqplyrge lnehlgllgp yiraevedni mvtfrnqasr
pysfyss1is yeedqrqgae prknfvkpne tktyfwkvqh hmaptkdefd ckawayfsdv
dlekdvhsgl igpllvchtn tlnpahgrqv tvqefalfft ifdetkswyf tenmerncra
pcniqmedpt fkenyrfhai ngyimdtlpg lvmaqdrir wyllsmgsne nihsihfsg1h
vftvrk1key kmalynlypg vftvemlps kagiwrvecl igehl1hagms t1flvysnkc
qtplgmasgh irdfqitasg qygqwapkla r1hysgsina wstkepfswi kvdl1lapmi
hgiktqgarq kfsslyisqf iimysldgkk wqtyrgnstg tlmvffgnvd ssgikhn1fn
ppiaryirl hpthysirst lrmelmgcdl nscsmplgme skaisdaqit assyftnmfa
twspskar1h lqgrsnawrp qvnnpkewlq vdfqktmkvt gvt1tqgvksl ltsmyvkefl
iss1qdghq1w t1ffqngkvk vfqgnqdsft pvvnsl1dppl ltrylrihpq swvhqialrm
evlgcea1dl y (SEQ ID NO:39)

```

Hinh 16

23/28

BDL004 - agc ttcagccaga atccacctgt cctgaaacgc caccagagg (SEQ ID NO:40)
BDLNG1 - agcttcagccagaaatGTGAGCAACAATGTGAGCAACAATGCCACCAATAATGTACCAACccacctgtcctgaaacgccaccagagg (SEQ ID NO:41)
BDLNG4 - agcttcagccagaaatGTGAGCAACAATGCCACCAACAATGTGAGCAACccacctgtcctgaaacgccaccagagg (SEQ ID NO:42)
BDLNG5 - agcttcagccagaaatGTGAGCAATAATGCCACCAACccacctgtcctgaaacgccaccagagg (SEQ ID NO:43)
BDLNG6 - agcttcagccagaaatGTGAGCAATAATccacctgtcctgaaacgccaccagagg (SEQ ID NO:44)
BDLNG9 - agcttcagccagaaatAGGAGCCTGccacctgtcctgaaacgccaccagagg (SEQ ID NO:45)
BDLNG10 - agcttcagccagaaatGCCACTAATGTGTCTAACAACTCTGCTACCTCTGCTGACTCTGTGAGCCccacctgtcctgaaacgccaccagagg (SEQ ID NO:46)
BDLNG16 - agcttcagccagaaatGCCACCAACTATGTGAACAGGAGCCTGccacctgtcctgaaacgccaccagagg (SEQ ID NO:47)
BDLNG17 - agcttcagccagaaatGCCACCAACTATGTGAACAGGAGCCTGTCTTGCCACCTCTGCTGACTCTGTGAGCCAGAAATccacctgtcctgaaacgccaccagagg (SEQ ID NO:48)
BDLNG18 - agcttcagccagaaatGTGAGCAACAATGTGAGCAATGTGTCTGTCTGCTTccacctgtcctgaaacgccaccagagg (SEQ ID NO:49)
BDLNG19 - agcttcagccagaaatATCACTGTGGCCTCTGCCACCTCTAACATCACTGTGGCCTCTGCTGACccacctgtcctgaaacgccaccagagg (SEQ ID NO:50)
BDLNG20 - agcttcagccagaaatCACTGTGACCAACATCACTGTGACTGCCccacctgtcctgaaacgccaccagagg (SEQ ID NO:51)
BDLNG21 - agcttcagccagaaatCAGACTGTGACCAACATCACTGTGACTGCCccacctgtcctgaaacgccaccagagg (SEQ ID NO:52)
BDLNGV - agcttcagccagaaatGCCACTAATGTGTCTAACACAGCAACACCAGCAATGACAGCAATGTGTCTTccacctgtcctgaaacgccaccagagg (SEQ ID NO:53)

Hình 17

24/28

Khung plasmid của CS12-CRM8.2-Vrp

gacgattttaaatgacaagcttggcgtaatcatggtcatagetgtttcctgtgtgaaattgttatccg
 ctcacaattccacacaacatacagagccggaagcataaagtgtaaagcctgggggtgcctaagtgtga
 gctaactcacattaattgcggttgcgtcactgcccgtttccagtcgggaaacctgtcgtgccagct
 gcattaatgaatcggccaacgcgcggggagaggcggtttgcgattgggcgctcttccgcttcctcg
 ctactgactcgtcgcgtcggtcgttcggctgcgggcagcggtatcagctcactcaaaggcggtaa
 tacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggc
 caggaaccgtaaaaaggccgcttgcgtggcggtttttccataggtccgccccctgacgagcatcac
 aaaaatcgagcgtcaagtcagaggtggcgaaaaccgcaggaactataaagataaccaggcggttcccc
 ctggaagctccctcgtcgcgtctcctgttccgaccctgccgcttaccggataacctgtccgcctttct
 cccttcgggaagcgtggcgcttttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcggt
 cgctccaagctgggctgtgtgcacgaaccccccggttcagcccagccgctgcgccttatccggtaact
 atcgtcttgagtcacaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggat
 tagcagagcagaggtatgtaggcggtgctacagagttccttgaagtggcctaactacggctacact
 agaagaacagtatgttggtatctgcgtctcgtgaagccagttaccttcggaaaaagagttggtagct
 cttgatccggcaacaaaccacgcgtggttagcggtgggtttttgtttgcaagcagcagattacgcg
 cagaaaaaaaggatctcaagaagatcctttgatcttttctacggggctcgtacgctcagtggaacgaa
 aactcacgttaagggattttgggtcatgagattatcaaaaaggatcttcacctagatccttttaatt
 aaaaatgaagttttaaatcaatctaaagtatatatgagtaaaacttggctcgtacagttaccaatgctt
 aatcagtgaggcacctatctcagcgatctgtctatttcgttcacatagttgcctgactccccgctc
 gtgtagataactacgatacgggagggccttaccatctggccccagtgctgcaatgataccgcgagacc
 cacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtgg
 tcctgcaactttatccgcctccatccagtcctattaattgttgccgggaagctagagtaagtagttcg
 ccagttaatagtttgcgcaacggttgttgccattgtctacaggcatcgtggtgtcacgctcgtcggttg
 gtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatccccatgttggtgcaa
 aaaagcgggttagctccttcgggtcctccgatcgttgtcagaagtaagttggccgcagtggtatcactc
 atgggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactg
 gtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcggcgctc
 aatacgggataaataaccgcgccacatagcagaactttaaaagtgtcatcattggaaaacgttcttcg
 gggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaaacccactcgtgcaccca
 actgatcttcagcatcttttactttcaccagcggttctgggtgagcaaaaacaggaaggcaaaatgc
 cgcaaaaaaggggaataaagggcgacacggaaatgttgaaatactcatactcttccctttttcaatattat
 tgaagcatttatcagggttattgtctcatgagcggatacatatttgaaatgtatttagaaaaataaac
 aaataggggttccgcgcacatttccccgaaaagtgccacctgacgtctaagaaaccattattatcat
 gacattaacctataaaaaataggcgatatcacgaggccctttcgtctcgcgcgttttcggtgatgacggt
 gaaaacctctgacacatgcagctcccgagacgggtcacagcttgtctgtaagcggatgccgggagca
 gacaagcccgtcagggcgctcagcgggtgttgccgggtgtcggggctggcttaactatgcggcatc
 agagcagattgtactgagagtgcaccatatgcgggtgtgaaataccgcacagatgcgtaaggagaaaa
 taccgcatecaggcgccattcgccattcaggctgcgcaactgttgggaagggcgatcgggtcgggcct
 cTtcgctattacgccagctggcgaaaggggatgtgctgcaaggcgattaagttgggtaacgccagg
 gttttccagtcacgacgttgtaaaacgacggccagtgaaattcctcgagatttaaatgacg

(SEQ ID NO:54)

Hình 18

25/28

Đơn vị sao chép pMB1

tggcgtaatc atgggtcatag ctgttttcctg tgtgaaattg ttatccgctc acaattccac
 acaacatacg agccggaagc ataaagtgtg aagcctgggg tgccataatga gtgagctaac
 tcacattaat tgcgttgcg cactgcccg ctttccagtc gggaaacctg tcgtgccagc
 tgcattaatg aatcggccaa cgcgcgggga gaggcgggtt gcgtattggg cgtcttccg
 cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc
 actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt
 gagcaaaaag ccagcaaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc
 ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa
 acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc
 ctgttccgac cctgccgctt accggatacc tgtcgcgctt tctcccttcg ggaagcgtgg
 cgcttttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcggt cgctccaagc
 tgggctgtgt gcacgaacc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc
 gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca
 ggattagcag agcgaggtat gtaggcgggt ctacagagtt cttgaagtgg tggcctaact
 acggctacac tagaagaaca gtatttggtg tctgcgctct gctgaagcca gttaccttcg
 gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtgggtttt
 ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat ccttt
 (SEQ ID NO:55)

Bla (ApR)

ttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat
 ccatagttgc ctgactcccc gtctgttaga taactacgat acgggagggc ttaccatctg
 gcccagtg c tgcattgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa
 taaaccagcc agccggaagg gccgagcgca gaagtgggtc tgcaacttta tccgcctcca
 tccagctctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc
 gcaacgttgt tgccattgct acaggcatcg tgggtgtcacg ctcgctcgtt ggtatggctt
 cattcagctc cggttcccaa cgatcaaggc gagttacatg atccccatg ttgtgcaaaa
 aagcggtag ctcttcgggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat
 cactcatggg tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct
 tttctgtgac tgggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga
 gttgctcttg cccggcgta atacgggata ataccgcgc acatagcaga actttaaaag
 tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga
 gatccagttc gatgtaaccc actcgtgcac ccaactgac ttcagcatct tttactttca
 ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg
 cgacacggaa atgttgaata ctcat (SEQ ID NO:56)

26/28

CS12-CRM8.2-Vp

tgcgcggtttcgggtgatgacgggtgaaaacctctgaacacatgcagctcccggagacgggtcacagcttg
tctgtaagcggatgccgggagcagacaagcccgtcagggcgcgctcagcgggtgttgccgggtgtcgg
ggctggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcgggtgtgaaatac
cgcacagatgcgtaaggagaaaaataccgcacatcaggcgccattcgccattcagggtgcgcaactgttg
ggaagggcgatcgggtcggggcctcttcgctattacgccagctggcgaaagggggatgtgctgcaagg
cgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgaattcc
tcgagatttaaatgacgttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgacc
aaaggtcgcgcgacgcccgggttttgcggggcgccctcagtgcgagcgagcgcgagagagggga
gtggccaactccatcactaggggttccctgagtttaaaccttcgtcgacggggaggctgctgggtgaat
attaaccaaggtcaccccagttatcggaggagcaaacaggggctaagtccaccggggaggctgctg
gtgaatattaaccaaggtcaccccagttatcggaggagcaaacaggggctaagtccacaaatgacct
attaagaatatttcatagaacgaatgttccgatgctctaattctcttagacaagggttcataatttgta
tgggttacttattctctctttgttgactaagtcataataatcagaatcagcaggtttgcagtcagattg
gcagggataagcagcctagctcaggagaagtgagtataaaagccccaggctgggagcagccatcaca
gaagtccactcattcttggcagggggcgcgcccgccaccatgcagattgagctgagcacctgcttcttc
ctgtgcctgctgaggttctgcttctctgccaccaggagatactacctgggggctgtggagcttctctt
gggactacatgcagctctgacctgggggagctgcctgtggatgccaggttcccaccagagtgcccaa
atccttcccattcaacacctctgtggtctacaagaagacctcttctgtggagttcactgaccacctg
ttcaacattgccaaaccaggccaccctggatgggactcctgggaccaccattcaggctgaggtgt
atgacactgtggtcgtcacccctcaagaacatggcctcccacctgtgagcctgcatgctgtgggggt
cagctactggaagtcctctgagggggctgagtatgatgaccagacctcccagagggagaaggaggat
gacaaagtgttccctgggaagagccacacctatgtgtggcaggtcctcaaggagaatggccccactg
cctctgaccaccctgcctgacctactcctaccttctcatgtggacctggtcaaggacctcaactc
tggaactgattggggccctgctgggtgtgcagggagggctccctggccaaagagaagaccagacctg
cacaagttcattctcctgtttgctgtctttgatgagggcaagagctggcactctgaaaccaagaact
ccctgatgcaggacagggtgctgcctctgccagggcctggcccaagatgcacactgtgaatggcta
tgtgaacaggagcctgcctggactcattggctgccacaggaaatctgtctactggcatgtgattggc
atggggacaacccctgaggtgcactccatttctcctggaggggccacaccttccctggtcaggaaccaca
gacaggccagcctggagatcagccccatcaccttccctcactgccagacctgctgatggacctcg
acagttcctgctgttctgccacatcagctcccaccagcatgatggcatggaggccctatgtcaaggtg
gacagctgccctgaggagccacagctcaggatgaagaacaatgaggaggctgaggactatgatgatg
acctgactgactctgagatggatgtggtccgctttgatgatgacaacagcccatccttcattcagat
caggtctgtggccaagaacacccccaaagacctgggtgcactacattgctgctgaggaggaggactgg
gactatgccccactggtcctggcccctgatgacaggagctacaagagccagtacctcaacaatggcc
cacagaggattggacgcaagtacaagaaagtcaggttcatggcctacactgatgaaaccttcaagac

(Tiếp)

Hình 20A

28/28

aaccaggacagcttcacccctgtggtgaacagcctggacccccccctcctgaccagatacctgagga
ttcacccccagagctgggtccaccagattgccctgaggatggaggtcctgggatgtgaggcccagga
cctgtaactgatgagcgccgctcttagtagcagtatcgataataaaaagatctttattttcattagat
ctgtgtgttggttttttgtgtgttaattaagctcgcggaaggaacccctagtgtgaggttgccact
ccctctctgcgcgctcgtcgcctcactgagggcgggcgaccaaaggtcgcccgacgcccgggctttg
cccgggcgccctcagtgtgagcgagcgagcgcgagagggagtgccaagacgatttaaatgacaag
cttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaac
atacgagccggaagcataaagtgtaaagcctgggggtgcctaatagtgtgagctaacacattaattg
cgttgcgctcactgcccgtttccagtcgggaaacotgtcgtgccagctgcattaatgaatcgcca
acgcgcggggagagggcggtttgcgtattgggegetcttccgcttccctcgtcactgactcgtgcgc
tcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggttaatacggttatccacagaat
caggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaaggc
cgcggtgttggtggtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagt
cagaggtggcgaaacccgacaggactataaagataccaggcggtttccccctggaagctccctcgtgc
gctctcctgttccgacctgcccgttacccgatacctgtccgcttctcccttcgggaagcgtggc
gctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcgctccaagctgggctgt
gtgcacgaacccccgttcagcccgaacgctgcgccttatccggttaactatcgtcttgagtccaacc
cggtaagacacgacttatcgccactggcagcagccactggtaacaggatttagcagagcgaggtatgt
aggcggtgctacagagttcttgaagtgggtggcctaactacggctacactagaagaacagtatttggt
atctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaa
ccaccgctggttagcgggtggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctca
agaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaaactcacgttaagggatt
ttggtcatgagattatcaaaaaaggatcttcacctagatccttttaattaaaaatgaagttttaaat
caatctaaagtatatatgagtaaaacttgggtctgacagttaccaatgcttaatcagtgtggcacctat
ctcagcgatctgtctattttcgttcacatcatagttgcctgactccccgctcgtgtagataactacgata
cgggagggttaccatctggccccagtgctgcaatgataccgcgagaccacgctcaccgggtccag
atttatcagcaataaaaccagccagccggaaggccgagcgcagaagtggctcgtgcaactttatccgc
ctccatccagctctattaattgtttgcccgggaagctagagtaagtagttcgcacagttaatagtttgcc
aacgttggttgccatttgctacaggcatcgtgggtgtcacgctcgtcgttttggtatggcttcattcagct
ccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcgggttagctcctt
cggctcctccgatcgttgtcagaagtaagttggccgcagtggttatcactcatggttatggcagcactg
cataattctcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtaactcaaccaagt
cattctgagaatagtgtatgcggcgacccaggttgctcttgcccggtcaatacgggataataccgc
gccacatagcagaactttaaaagtgtcatcattggaaaacggttcttcggggcgaaaaactctcaagg
atcttacgcgtgttgagatccagttcgatgtaaccactcgtgcacccaactgatcttcagcatctt
ttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaaagggaataag
ggcgacacggaaatgttgaatactcactcttctttttcaatattattgaagcatttatcaggggt
tattgtctcatgagcgatacatatttgatgtatttagaaaaataaacaataaggggttccgcgca
catttccccgaaaagtgccacctgacgtctaagaaaccattattatcatgacattaacctataaaaa
taggcgtatcacgaggccctttcgtc (SEQ ID NO:57)

Hinh 20C

DANH MỤC TRÌNH TỰ

<110> BAXALTA INCORPORATED
BAXALTA GMBH

<120> HỢP PHẦN AXIT NUCLEIC CÓ CHỨA POLYNUCLEOTIT YẾU TỐ VIII VÀ VECTO LIỆU
PHÁP GEN ĐỘNG VẬT CÓ VÚ CÓ CHỨA NÓ

<130> 008073-5191-WO

<140> PCT/US2020/013722

<141> 15/01/2020

<150> US 62/793058

<151> 16/01/2019

<160> 57

<170> PatentIn phiên bản 3.5

<210> 1

<211> 4392

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tổng hợp

<400> 1

atgcagattg agctgagcac ctgcttcttc ctgtgcctgc tgaggttctg cttctctgcc	60
accaggagat actacctggg ggctgtggag ctttcttggg actacatgca gtctgacctg	120
ggggagctgc ctgtggatgc caggttccca cccagagtgc ccaaatacctt cccattcaac	180
acctctgtgg tctacaagaa gaccctcttt gtggagttca ctgaccacct gttcaacatt	240
gccaaaccca ggccaccctg gatgggactc ctgggaccca ccattcaggc tgagggtgat	300
gacactgtgg tcgtcacct caagaacatg gcctcccacc ctgtgagcct gcatgctgtg	360
ggggtcagct actggaagtc ctctgagggg gctgagtatg atgaccagac ctcccagagg	420
gagaaggagg atgacaaaagt gttccctggg aagagccaca cctatgtgtg gcaggctctc	480
aaggagaatg gcccactgc ctctgaccca ccctgcctga cctactccta cttttctcat	540
gtggacctgg tcaaggacct caactctgga ctgattgggg ccctgctggt gtgcaggag	600
ggctccctgg ccaaagagaa gaccagacc ctgcacaagt tcattctcct gtttgctgtc	660
tttgatgagg gcaagagctg gcactctgaa accaagaact ccctgatgca ggacagggat	720
gctgcctctg ccagggcctg gcccaagatg cacactgtga atggctatgt gaacaggagc	780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg	840

acaaaccttg aggtgcactc cattttcctg gagggccaca ccttcctggt caggaaccac 900
 agacaggcca gcctggagat cagcccatc accttcctca ctgccagac cctgctgatg 960
 gacctcggac agttcctgct gttctgccac atcagctccc accagcatga tggcatggag 1020
 gcctatgtca aggtggacag ctgccctgag gagccacagc tcaggatgaa gaacaatgag 1080
 gaggtgagg actatgatga tgacctgact gactctgaga tggatgtggt ccgctttgat 1140
 gatgacaaca gcccatcctt cattcagatc aggtctgtgg ccaagaaaca cccaagacc 1200
 tgggtgcact acattgctgc tgaggaggag gactgggact atgccccact ggtcctggcc 1260
 cctgatgaca ggagctacaa gagccagtac ctcaacaatg gccacagag gattggacgc 1320
 aagtacaaga aagtcagggt catggcctac actgatgaaa ccttcaagac cagggaggcc 1380
 attcagcatg agtctggcat cctgggcca ctctgtatg gggagggtgg ggacaccctg 1440
 ctcatcatct tcaagaacca ggctccagg ccctacaaca tctaccaca tggcatcact 1500
 gatgtcaggc cctgtacag ccgcaggctg ccaaagggg tgaaacacct caaggacttc 1560
 cccattctgc ctggggagat cttcaagtac aagtggactg tcaactgtga ggatggacca 1620
 accaaatctg accccagggt cctcaccaga tactactcca gctttgtgaa catggagagg 1680
 gacctggcct ctggcctgat tggccactg ctcatctgct acaaggagtc tgtggaccag 1740
 aggggaaacc agatcatgtc tgacaagagg aatgtgattc tgttctctgt ctttgatgag 1800
 aacaggagct ggtacctgac tgagaacatt cagcgcttcc tgccaaccc tgctggggtg 1860
 cagctggagg accctgagtt ccaggccagc aacatcatgc actccatcaa tggctatgtg 1920
 tttgacagcc tccagcttct tgtctgctg catgagggtg cctactggta cattcttct 1980
 attggggccc agactgactt ctttctgtc ttcttctctg gctacacctt caaacacaag 2040
 atggtgtatg aggacaccct gacctcttc ccattctctg gggagactgt gttcatgagc 2100
 atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc 2160
 atgactgcc tgctcaaagt ctctcctgt gacaagaaca ctggggacta ctatgaggac 2220
 agctatgagg acatctctgc ctacctgctc agcaagaaca atgccattga gccaggagc 2280
 ttcagccaga atgtgagcaa taatgccacc aaccacctg tctgaaacg ccaccagagg 2340
 gagatcacca ggaccaccct ccagtctgac caggaggaga ttgactatga tgacaccatt 2400
 tctgtggaga tgaagaaaga ggactttgac atctatgacg aggacgagaa ccagagccca 2460
 aggagcttcc agaagaagac caggcactac ttcattgctg ctgtggagcg cctgtgggac 2520

tatggcatga gctccagccc ccatgtcctc aggaacaggg cccagtctgg ctctgtgcca	2580
cagttcaaga aagtggctctt ccaagagttc actgatggca gcttcaccca gccoctgtac	2640
agaggggagc tgaatgagca cctgggactc ctgggcccac acatcagggc tgagggtggag	2700
gacaacatca tgggtgacctt ccgcaaccag gcctccaggc cctacagctt ctacagctcc	2760
ctcatcagct atgaggagga ccagaggcag ggggctgagc cacgcaagaa ctttgtgaaa	2820
cccaatgaaa ccaagacctt cttctggaaa gtccagcacc acatggcccc caccaaggat	2880
gagtttgact gcaaggcctg ggccacttct tctgatgtgg acctggagaa ggatgtgcac	2940
tctggcctga ttggcccact cctgggtctgc cacaccaaca ccctgaacct tgcccatgga	3000
aggcaagtga ctgtgcagga gtttgccctc ttcttcacca tctttgatga aaccaagagc	3060
tggctacttca ctgagaacat ggagcgcaac tgcagggccc catgcaacat tcagatggag	3120
gaccccacct tcaaagagaa ctaccgcttc catgccatca atggctacat catggacacc	3180
ctgcctgggc ttgtcatggc ccaggaccag aggatcaggt ggtacctgct ttctatgggc	3240
tccaatgaga acattcactc catccacttc tctgggcatg tcttctactgt gcgcaagaag	3300
gaggagtaca agatggccct gtacaacctc taccctgggg tctttgagac tgtggagatg	3360
ctgccctcca aagctggcat ctggagggtg gactgcctca ttggggagca cctgcatgct	3420
ggcatgagca ccctgttctt ggtctacagc aacaagtgcc agacccccct gggaatggcc	3480
tctggccaca tcagggactt ccagatcact gcctctggcc agtatggcca gtgggcccc	3540
aagctggcca ggctccacta ctctggatcc atcaatgctt ggagcaccaa ggagccattc	3600
agctggatca aagtggacct gctggcccc atgatcatcc atggcatcaa gacccagggg	3660
gccaggcaga agttctccag cctgtacatc agccagttca tcatcatgta cagcctggat	3720
ggcaagaaat ggcagacctt cagaggcaac tccactggaa cactcatggt cttctttggc	3780
aatgtggaca gctctggcat caagcacaac atcttcaacc cccaatcat cgccagatac	3840
atcaggctgc accccaccca ctacagcatc cgcagcacc tcaggatgga gctgatgggc	3900
tgtgacctga actcctgcag catgcccctg ggcatggaga gcaaggccat ttctgatgcc	3960
cagatcactg cctccagcta cttcaccaac atgtttgcc cctggagccc aagcaaggcc	4020
aggctgcacc tccagggaag gagcaatgcc tggaggcccc aggtcaacaa cccaaaggag	4080
tggctgcagg tggacttcca gaagaccatg aaggtcactg gggtgaccac ccagggggtc	4140
aagagcctgc tcaccagcat gtatgtgaag gagttcctga tcagctccag ccaggatggc	4200

caccagtgga ccctcttctt ccagaatggc aaggtcaagg tgttccaggg caaccaggac 4260
 agcttcaccc ctgtggtgaa cagcctggac cccccctcc tgaccagata cctgaggatt 4320
 cccccccaga gctgggtcca ccagattgcc ctgaggatgg aggtcctggg atgtgaggcc 4380
 caggacctgt ac 4392

<210> 2
 <211> 1470
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 2

Cys Ser Phe Leu Ala Ala Met Gln Ile Glu Leu Ser Thr Cys Phe Phe
 1 5 10 15

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

Gly Ala Val Glu Leu Ser Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu
 35 40 45

Leu Pro Val Asp Ala Arg Phe Pro Pro Arg Val Pro Lys Ser Phe Pro
 50 55 60

Phe Asn Thr Ser Val Val Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr
 65 70 75 80

Asp His Leu Phe Asn Ile Ala Lys Pro Arg Pro Pro Trp Met Gly Leu
 85 90 95

Leu Gly Pro Thr Ile Gln Ala Glu Val Tyr Asp Thr Val Val Val Thr
 100 105 110

Leu Lys Asn Met Ala Ser His Pro Val Ser Leu His Ala Val Gly Val
 115 120 125

Ser Tyr Trp Lys Ser Ser Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser
 130 135 140

Gln Arg Glu Lys Glu Asp Asp Lys Val Phe Pro Gly Lys Ser His Thr
 145 150 155 160

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

Pro Cys Leu Thr Tyr Ser Tyr Leu Ser His Val Asp Leu Val Lys Asp
 180 185 190

Leu Asn Ser Gly Leu Ile Gly Ala Leu Leu Val Cys Arg Glu Gly Ser
 195 200 205

Leu Ala Lys Glu Lys Thr Gln Thr Leu His Lys Phe Ile Leu Leu Phe
 210 215 220

Ala Val Phe Asp Glu Gly Lys Ser Trp His Ser Glu Thr Lys Asn Ser
 225 230 235 240

Leu Met Gln Asp Arg Asp Ala Ala Ser Ala Arg Ala Trp Pro Lys Met
 245 250 255

His Thr Val Asn Gly Tyr Val Asn Arg Ser Leu Pro Gly Leu Ile Gly
 260 265 270

Cys His Arg Lys Ser Val Tyr Trp His Val Ile Gly Met Gly Thr Thr
 275 280 285

Pro Glu Val His Ser Ile Phe Leu Glu Gly His Thr Phe Leu Val Arg
 290 295 300

Asn His Arg Gln Ala Ser Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr
 305 310 315 320

Ala Gln Thr Leu Leu Met Asp Leu Gly Gln Phe Leu Leu Phe Cys His
 325 330 335

Ile Ser Ser His Gln His Asp Gly Met Glu Ala Tyr Val Lys Val Asp
 340 345 350

Ser Cys Pro Glu Glu Pro Gln Leu Arg Met Lys Asn Asn Glu Glu Ala
 355 360 365

Glu Asp Tyr Asp Asp Asp Leu Thr Asp Ser Glu Met Asp Val Val Arg
370 375 380

Phe Asp Asp Asp Asn Ser Pro Ser Phe Ile Gln Ile Arg Ser Val Ala
385 390 395 400

Lys Lys His Pro Lys Thr Trp Val His Tyr Ile Ala Ala Glu Glu Glu
405 410 415

Asp Trp Asp Tyr Ala Pro Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr
420 425 430

Lys Ser Gln Tyr Leu Asn Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr
435 440 445

Lys Lys Val Arg Phe Met Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg
450 455 460

Glu Ala Ile Gln His Glu Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly
465 470 475 480

Glu Val Gly Asp Thr Leu Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg
485 490 495

Pro Tyr Asn Ile Tyr Pro His Gly Ile Thr Asp Val Arg Pro Leu Tyr
500 505 510

Ser Arg Arg Leu Pro Lys Gly Val Lys His Leu Lys Asp Phe Pro Ile
515 520 525

Leu Pro Gly Glu Ile Phe Lys Tyr Lys Trp Thr Val Thr Val Glu Asp
530 535 540

Gly Pro Thr Lys Ser Asp Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser
545 550 555 560

Phe Val Asn Met Glu Arg Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu
565 570 575

Leu Ile Cys Tyr Lys Glu Ser Val Asp Gln Arg Gly Asn Gln Ile Met
580 585 590

Ser Asp Lys Arg Asn Val Ile Leu Phe Ser Val Phe Asp Glu Asn Arg
595 600 605

Ser Trp Tyr Leu Thr Glu Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala
610 615 620

Gly Val Gln Leu Glu Asp Pro Glu Phe Gln Ala Ser Asn Ile Met His
625 630 635 640

Ser Ile Asn Gly Tyr Val Phe Asp Ser Leu Gln Leu Ser Val Cys Leu
645 650 655

His Glu Val Ala Tyr Trp Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp
660 665 670

Phe Leu Ser Val Phe Phe Ser Gly Tyr Thr Phe Lys His Lys Met Val
675 680 685

Tyr Glu Asp Thr Leu Thr Leu Phe Pro Phe Ser Gly Glu Thr Val Phe
690 695 700

Met Ser Met Glu Asn Pro Gly Leu Trp Ile Leu Gly Cys His Asn Ser
705 710 715 720

Asp Phe Arg Asn Arg Gly Met Thr Ala Leu Leu Lys Val Ser Ser Cys
725 730 735

Asp Lys Asn Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser
740 745 750

Ala Tyr Leu Leu Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser
755 760 765

Gln Asn Val Ser Asn Asn Ala Thr Asn Pro Pro Val Leu Lys Arg His
770 775 780

Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile
785 790 795 800

Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp
805 810 815

Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys
820 825 830

Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly
835 840 845

Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser
850 855 860

Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly Ser
865 870 875 880

Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly Leu
885 890 895

Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr
900 905 910

Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile
915 920 925

Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe
930 935 940

Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His His
945 950 955 960

Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe
965 970 975

Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly Pro
980 985 990

Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln
995 1000 1005

Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu
1010 1015 1020

Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg
1025 1030 1035

Ala	Pro	Cys	Asn	Ile	Gln	Met	Glu	Asp	Pro	Thr	Phe	Lys	Glu	Asn
1040						1045					1050			
Tyr	Arg	Phe	His	Ala	Ile	Asn	Gly	Tyr	Ile	Met	Asp	Thr	Leu	Pro
1055						1060					1065			
Gly	Leu	Val	Met	Ala	Gln	Asp	Gln	Arg	Ile	Arg	Trp	Tyr	Leu	Leu
1070						1075					1080			
Ser	Met	Gly	Ser	Asn	Glu	Asn	Ile	His	Ser	Ile	His	Phe	Ser	Gly
1085						1090					1095			
His	Val	Phe	Thr	Val	Arg	Lys	Lys	Glu	Glu	Tyr	Lys	Met	Ala	Leu
1100						1105					1110			
Tyr	Asn	Leu	Tyr	Pro	Gly	Val	Phe	Glu	Thr	Val	Glu	Met	Leu	Pro
1115						1120					1125			
Ser	Lys	Ala	Gly	Ile	Trp	Arg	Val	Glu	Cys	Leu	Ile	Gly	Glu	His
1130						1135					1140			
Leu	His	Ala	Gly	Met	Ser	Thr	Leu	Phe	Leu	Val	Tyr	Ser	Asn	Lys
1145						1150					1155			
Cys	Gln	Thr	Pro	Leu	Gly	Met	Ala	Ser	Gly	His	Ile	Arg	Asp	Phe
1160						1165					1170			
Gln	Ile	Thr	Ala	Ser	Gly	Gln	Tyr	Gly	Gln	Trp	Ala	Pro	Lys	Leu
1175						1180					1185			
Ala	Arg	Leu	His	Tyr	Ser	Gly	Ser	Ile	Asn	Ala	Trp	Ser	Thr	Lys
1190						1195					1200			
Glu	Pro	Phe	Ser	Trp	Ile	Lys	Val	Asp	Leu	Leu	Ala	Pro	Met	Ile
1205						1210					1215			
Ile	His	Gly	Ile	Lys	Thr	Gln	Gly	Ala	Arg	Gln	Lys	Phe	Ser	Ser
1220						1225					1230			
Leu	Tyr	Ile	Ser	Gln	Phe	Ile	Ile	Met	Tyr	Ser	Leu	Asp	Gly	Lys
1235						1240					1245			

Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val
 1250 1255 1260
 Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe
 1265 1270 1275
 Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His
 1280 1285 1290
 Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp
 1295 1300 1305
 Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile
 1310 1315 1320
 Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe
 1325 1330 1335
 Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg
 1340 1345 1350
 Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu
 1355 1360 1365
 Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr
 1370 1375 1380
 Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe
 1385 1390 1395
 Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe
 1400 1405 1410
 Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe
 1415 1420 1425
 Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr
 1430 1435 1440
 Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg
 1445 1450 1455

Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr
 1460 1465 1470

<210> 3
 <211> 5120
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 3
 ttggccactc cctctctgcg cgctcgctcg ctcaactgagg ccgggcgacc aaaggctcgcc 60
 cgacgccccg gctttgcccc ggcgggcctca gtgagcgagc gagcgcgag agaggggagtg 120
 gccaaactcca tcactagggg ttcttggggg aggctgctgg tgaatattaa ccaaggtcac 180
 ccagttatc ggaggagcaa acaggggcta agtccaccgg gggaggctgc tggatgaatat 240
 taaccaaggt caccacagtt atcgaggag caaacagggg ctaagtccac aaatgacct 300
 ttaagaatat ttcatagaac gaatgttccg atgctctaata ctctctagac aagggttcata 360
 tttgtatggg ttacttattc tctctttgtt gactaagtca ataatacagaa tcagcaggtt 420
 tgcagtcaga ttggcaggga taagcagcct agctcaggag aagtgagtat aaaagcccca 480
 ggctgggagc agccatcaca gaagtccact cattcttggc aggcaccat gcagattgag 540
 ctgagcacct gcttcttctt gtgcctgctg aggttctgct tctctgccac caggagatac 600
 tacctggggg ctgtggagct ttcttgggac tacatgcagt ctgacctggg ggagctgcct 660
 gtggatgcca ggttcccacc cagagtgcac aaatccttcc cattcaacac ctctgtggtc 720
 tacaagaaga cctcttttgt ggagttcact gaccacctgt tcaacattgc caaaccagg 780
 ccacctgga tgggactcct gggaccacc attcaggctg aggtgtatga cactgtggtc 840
 gtcacctca agaacatggc ctcccacct gtgagcctgc atgctgtggg ggtcagctac 900
 tggaagtcct ctgagggggc tgagtatgat gaccagacct ccagaggga gaaggaggat 960
 gacaaagtgt tccctgggaa gagccacacc tatgtgtggc aggtcctcaa ggagaatggc 1020
 cccactgcct ctgaccacc ctgcctgacc tactcctacc tttctcatgt ggacctggtc 1080
 aaggacctca actctggact gattggggcc ctgctggtgt gcagggaggg ctccctggcc 1140
 aaagagaaga ccagaccct gcacaagtgc attctcctgt ttgctgtctt tgatgagggc 1200
 aagagctggc actctgaaac caagaactcc ctgatgcagg acagggatgc tgctctgccc 1260
 agggcctggc ccaagatgca cactgtgaat ggctatgtga acaggagcct gcctggactc 1320

attggctgcc acaggaaatc tgtctactgg catgtgattg gcatggggac aaccocctgag 1380
 gtgcactcca ttttcttgga gggccacacc ttcttggtca ggaaccacag acaggccagc 1440
 ctggagatca gcccacacac ctctctact gcccagaccc tgctgatgga cctcggacag 1500
 ttcttgctgt tctgccacat cagctccac cagcatgatg gcatggaggc ctatgtcaag 1560
 gtggacagct gccctgagga gccacagctc aggatgaaga acaatgagga ggctgaggac 1620
 tatgatgatg acctgactga ctctgagatg gatgtggtcc gctttgatga tgacaacagc 1680
 ccaccttca ttcagatcag gtctgtggcc aagaaacacc ccaagacctg ggtgcactac 1740
 attgctgctg aggaggagga ctgggactat gcccactgg tcctggcccc tgatgacagg 1800
 agctacaaga gccagtacct caacaatggc ccacagagga ttggacgcaa gtacaagaaa 1860
 gtcaggttca tggcctacac tgatgaaacc ttcaagacca gggaggccat tcagcatgag 1920
 tctggcatcc tgggcccact cctgtatggg gaggtggggg acaccctgct catcatcttc 1980
 aagaaccagg cctccaggcc ctacaacatc taccacatg gcatcactga tgtcaggccc 2040
 ctgtacagcc gcaggctgcc aaaggggggtg aaacacctca aggacttccc cattctgcct 2100
 ggggagatct tcaagtacaa gtggactgtc actgtggagg atggaccaac caaatctgac 2160
 ccaggtgcc tcaccagata ctactccagc tttgtgaaca tggagaggga cctggcctct 2220
 ggctgattg gccactgct catctgctac aaggagtctg tggaccagag gggaaaccag 2280
 atcatgtctg acaagaggaa tgtgattctg ttctctgtct ttgatgagaa caggagctgg 2340
 tacctgactg agaacattca gcgcttctg cccaacctg ctggggtgca gctggaggac 2400
 cctgagttcc aggccagcaa catcatgcac tccatcaatg gctatgtgtt tgacagcctc 2460
 cagctttctg tctgcctgca tgagggtggc tactggtaca ttctttctat tggggcccag 2520
 actgacttcc tttctgtctt ctctctggc tacaccttca aacacaagat ggtgtatgag 2580
 gacaccctga ccctcttccc attctctggg gagactgtgt tcatgagcat ggagaacct 2640
 ggctgtgga ttctgggatg ccacaactct gacttccgca acaggggcat gactgcctg 2700
 ctcaaagtct cctcctgtga caagaacact ggggactact atgaggacag ctatgaggac 2760
 atctctgcct acctgctcag caagaacaat gccattgagc ccaggagctt cagccagaat 2820
 gtgagcaata atgccaccaa cccacctgtc ctgaaacgcc accagaggga gatcaccagg 2880
 accaccctcc agtctgacca ggaggagatt gactatgatg acaccatttc tgtggagatg 2940
 aagaaagagg actttgacat ctatgacgag gacgagaacc agagcccaag gagcttccag 3000

aagaagacca ggcactactt cattgctgct gtggagcgcc tgtgggacta tggcatgagc 3060
tccagccccc atgtcctcag gaacagggcc cagtctggct ctgtgccaca gttcaagaaa 3120
gtggtcttcc aagagttcac tgatggcagc ttcacccagc cctgttacag aggggagctg 3180
aatgagcacc tgggactcct gggcccatac atcagggctg aggtggagga caacatcatg 3240
gtgaccttcc gcaaccaggc ctccaggccc tacagcttct acagctccct catcagctat 3300
gaggaggacc agaggcaggg ggctgagcca cgcaagaact ttgtgaaacc caatgaaacc 3360
aagacctact tctggaaagt ccagcaccac atggccccc ccaaggatga gtttgactgc 3420
aaggcctggg cctacttctc tgatgtggac ctggagaagg atgtgcactc tggcctgatt 3480
ggcccaactcc tggctctgcca caccaacacc ctgaacctg cccatggaag gcaagtgact 3540
gtgcaggagt ttgccctctt cttcaccatc tttgatgaaa ccaagagctg gtacttcact 3600
gagaacatgg agcgcaactg cagggcccca tgcaacattc agatggagga cccacacctc 3660
aaagagaact accgcttcca tgccatcaat ggctacatca tggacaccct gcctgggctt 3720
gtcatggccc aggaccagag gatcagggtg tacctgcttt ctatgggctc caatgagaac 3780
attcaactcca tccacttctc tgggcatgtc ttcactgtgc gcaagaagga ggagtacaag 3840
atggccctgt acaacctcta ccctggggctc tttgagactg tggagatgct gccctccaaa 3900
gctggcatct ggaggggtga gtgcctcatt ggggagcacc tgcatgctgg catgagcacc 3960
ctgttctctg tctacagcaa caagtgccag acccccctgg gaatggcctc tggccacatc 4020
agggaacttc agatcactgc ctctggccag tatggccagt gggcccccaa gctggccagg 4080
ctccactact ctggatccat caatgcctgg agcaccaagg agccattcag ctggatcaaa 4140
gtggacctgc tggcccccat gatcatccat ggcatcaaga cccagggggc caggcagaag 4200
ttctccagcc tgtacatcag ccagttcacc atcatgtaca gcctggatgg caagaaatgg 4260
cagacctaca gaggcaactc cactggaaca ctcatggtct tctttggcaa tgtggacagc 4320
tctggcatca agcacaacat cttcaacccc ccaatcatcg ccagatacat caggctgcac 4380
cccacccact acagcatccg cagcaccctc aggatggagc tgatgggctg tgacctgaac 4440
tctgcagca tgcccctggg catggagagc aaggccattt ctgatgccca gatcactgcc 4500
tccagctact tcaccaacat gtttgccacc tggagcccaa gcaaggccag gctgcacctc 4560
caggaagga gcaatgcctg gaggccccag gtcaacaacc caaaggagtg gctgcaggtg 4620
gacttccaga agaccatgaa ggtcactggg gtgaccaccc aggggggtcaa gagcctgctc 4680

accagcatgt atgtgaagga gttcctgate agctccagcc aggatggcca ccagtggacc 4740
 ctctttcttcc agaattggcaa ggtcaaggtg ttccagggca accaggacag ctccacccct 4800
 gtggtgaaca gcctggaccc cccctcctctg accagatacc tgaggattca cccccagagc 4860
 tgggtccacc agattgccct gaggatggag gtccctgggat gtgaggccca ggacctgtac 4920
 tgatgaaata aaagatcttt attttcatta gatctgtgtg ttggtttttt gtgtgaggaa 4980
 cccctagtga tggagttggc cactccctct ctgcgcgtc gctcgtcac tgaggccggg 5040
 cgaccaaagg tcgcccgcg cccgggcttt gccggggcg cctcagttag cgagcgagcg 5100
 cgagagagg gagtggccaa 5120

<210> 4
 <211> 145
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 4
 ttggccactc cctctctgcg cgctcgtcg ctactgagg ccggggcgacc aaaggtcgcc 60
 cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg 120
 gccaaactca tcactagggg ttctt 145

<210> 5
 <211> 72
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 5
 gggggaggct gctggtgaat attaaccaag gtcaccccag ttatcgagg agcaaacagg 60
 ggctaagtcc ac 72

<210> 6
 <211> 233
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 6
 aaatgaccta ttaagaatat ttcatagaac gaatgttccg atgctctaata ctctctagac 60
 aaggttcata tttgtatggg ttacttattc tctctttggt gactaagtca ataatacagaa 120
 tcagcaggtt tgcagtcaga ttggcaggga taagcagcct agctcaggag aagtgagtat 180
 aaaagcccca ggctggggagc agccatcaca gaagtccact cattcttggc agg 233

<210> 7
 <211> 5
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 7
 ccacc 5

<210> 8
 <211> 49
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 8
 aataaaagat ctttatcttc attagatctg tgtgttggtt ttttgtgtg 49

<210> 9
 <211> 145
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 9
 aggaaccctt agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcaactgagg 60
 ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggccggcctca gtgagcgagc 120
 gagcgcgag agagggagtg gccaa 145

<210> 10
 <211> 7794
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Tổng hợp

<400> 10

ttggccactc cctctctgcg cgctcgctcg ctactgagg cggggcgacc aaaggtcgcc	60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg	120
gccaactcca tctactagggg ttctctggggg aggtctgttg tgaatattaa ccaaggtcac	180
cccagttatc ggaggagcaa acaggggcta agtccaccgg gggaggctgc tggatgaatat	240
taaccaaggt cccccagtt atcgaggag caaacagggg ctaagtccac aaatgacct	300
ttaagaatat ttcatagaac gaatgttccg atgctctaata ctctctagac aaggttcata	360
tttgtatggg ttacttattc tctctttgtt gactaagtca ataatacagaa tcagcaggtt	420
tgcagtcaga ttggcaggga taagcagcct agctcaggag aagtgagtat aaaagcccca	480
ggctgggagc agccatcaca gaagtccact cattcttggc aggccaccat gcagattgag	540
ctgagcacct gcttcttctt gtgcctgctg aggttctgct tctctgccac caggagatac	600
tacctggggg ctgtggagct ttcttgggac tacatgcagt ctgacctggg ggagctgctt	660
gtggatgcca ggttcccacc cagagtgcac aaatccttcc cattcaacac ctctgtggtc	720
tacaagaaga cctcttttgt ggagttcact gaccacctgt tcaacattgc caaacccagg	780
ccacctgga tgggactcct gggaccacac attcaggctg aggtgtatga cactgtggtc	840
gtcacctca agaacatggc ctcccacct gtgagcctgc atgctgtggg ggtcagctac	900
tgaagtcct ctgagggggc tgagtatgat gaccagacct cccagaggga gaaggaggat	960
gacaaagtgt tccctgggaa gagccacacc tatgtgtggc aggtcctcaa ggagaatggc	1020
cccactgctt ctgaccacac ctgcctgacc tactcctacc tttctcatgt ggacctggtc	1080
aaggacctca actctggact gattggggcc ctgctggtgt gcagggaggg ctccctggcc	1140
aaagagaaga cccagacct gcacaagttc atttctctgt ttgctgtctt tgatgagggc	1200
aagagctggc actctgaaac caagaactcc ctgatgcagg acagggatgc tgctctgcc	1260
agggctggc ccaagatgca cactgtgaat ggctatgtga acaggagcct gcctggactc	1320
attggctgcc acaggaaatc tgtctactgg catgtgattg gcatggggac aacctctgag	1380
gtgcactcca ttttctgga gggccacacc ttctgggtca ggaaccacag acaggccagc	1440
ctggagatca gcccatacac ctctctact gccagaccc tgctgatgga cctcggacag	1500
ttctgtgtgt tctgccacat cagctccac cagcatgatg gcatggaggc ctatgtcaag	1560
gtggacagct gccctgagga gccacagctc aggatgaaga acaatgagga ggctgaggac	1620

tatgatgatg acctgactga ctctgagatg gatgtggtec gctttgatga tgacaacagc	1680
ccatccttca ttcagatcag gtctgtggcc aagaaacacc ccaagacctg ggtgcactac	1740
attgctgctg aggaggagga ctgggactat gccccactgg tccctggcccc tgatgacagg	1800
agctacaaga gccagtacct caacaatggc ccacagagga ttggacgcaa gtacaagaaa	1860
gtcagggttca tggcctacac tgatgaaacc ttcaagacca gggaggccat tcagcatgag	1920
tctggcatcc tgggcccact cctgtatggg gaggtggggg acaccttgct catcatcttc	1980
aagaaccagg cctccaggcc ctacaacatc taccacatg gcatcactga tgtcaggccc	2040
ctgtacagcc gcaggctgcc aaagggggtg aaacacctca aggacttccc cattctgcct	2100
ggggagatct tcaagtacaa gtggactgtc actgtggagg atggaccaac caaatctgac	2160
cccagggtgcc tcaccagata ctactccagc tttgtgaaca tggagaggga cctggcctct	2220
ggcctgattg gcccaactgct catctgctac aaggagtctg tggaccagag gggaaaccag	2280
atcatgtctg acaagaggaa tgtgattctg ttctctgtct ttgatgagaa caggagctgg	2340
tacctgactg agaacattca gcgcttcctg cccaacctg ctgggggtgca gctggaggac	2400
cctgagttcc aggccagcaa catcatgcac tccatcaatg gctatgtgtt tgacagcctc	2460
cagctttctg tctgcctgca tgagggtggc tactggtaca ttctttctat tggggcccag	2520
actgacttcc tttctgtctt cttctctggc tacaccttca aacacaagat ggtgtatgag	2580
gacacctga cctcttccc attctctggg gagactgtgt tcatgagcat ggagaacct	2640
ggcctgtgga ttctgggatg ccacaactct gacttccgca acaggggcat gactgccctg	2700
ctcaaagtct cctcctgtga caagaacact ggggactact atgaggacag ctatgaggac	2760
atctctgcct acctgctcag caagaacaat gccattgagc ccaggagctt cagccagaat	2820
gtgagcaata atgccaccaa cccacctgtc ctgaaacgcc accagaggga gatcaccagg	2880
accacctcc agtctgacca ggaggagatt gactatgatg acaccatttc tgtggagatg	2940
aagaaagagg actttgacat ctatgacgag gacgagaacc agagcccaag gagcttccag	3000
aagaagacca ggcactactt cattgctgct gtggagcgcc tgtgggacta tggcatgagc	3060
tccagcccc atgtcctcag gaacagggcc cagtctggct ctgtgccaca gttcaagaaa	3120
gtgggtcttcc aagagttcac tgatggcagc ttcacccagc cctgtacag aggggagctg	3180
aatgagcacc tgggactcct gggcccatac atcagggctg aggtggagga caacatcatg	3240
gtgaccttcc gcaaccaggc ctccaggccc tacagcttct acagctccct catcagctat	3300

gagggaggacc agaggcaggg ggctgagcca cgcaagaact ttgtgaaacc caatgaaacc 3360
 aagacctact tctggaaagt ccagcaccac atggccccca ccaaggatga gtttgactgc 3420
 aaggcctggg cctactttctc tgatgtggac ctggagaagg atgtgcactc tggcctgatt 3480
 ggccccactcc tggctctgcca caccaacacc ctgaaccctg cccatggaag gcaagtgact 3540
 gtgcaggagt ttgccctctt cttcaccatc tttgatgaaa ccaagagctg gtacttcact 3600
 gagaacatgg agcgcaactg caggggccca tgcaacattc agatggagga cccaccttc 3660
 aaagagaact accgcttcca tgccatcaat ggctacatca tggacaccct gcctgggctt 3720
 gtcattggccc aggaccagag gatcagggtg tacctgcttt ctatgggctc caatgagaac 3780
 attcactcca tccactttctc tgggcatgtc ttcactgtgc gcaagaagga ggagtacaag 3840
 atggccctgt acaacctcta ccctggggtc tttgagactg tggagatgct gccctccaaa 3900
 gctggcatct ggagggtgga gtgcctcatt ggggagcacc tgcattgctg catgagcacc 3960
 ctgttcctgg tctacagcaa caagtgccag acccccctgg gaatggcctc tggccacatc 4020
 agggacttcc agatcactgc ctctggccag tatggccagt gggcccccaa gctggccagg 4080
 ctccactact ctggatccat caatgcctgg agcaccaagg agccattcag ctggatcaaa 4140
 gtggacctgc tggcccccat gatcatccat ggcattcaaga cccagggggc caggcagaag 4200
 ttctccagcc tgtacatcag ccagttcctc atcatgtaca gcctggatgg caagaaatgg 4260
 cagacctaca gaggcaactc cactggaaca ctcatggtct tctttggcaa tgtggacagc 4320
 tctggcatca agcacaacat cttcaacccc ccaatcatcg ccagatacat caggctgcac 4380
 cccacccact acagcatccg cagcaccctc aggatggagc tgatgggctg tgacctgaac 4440
 tcttgacgca tggccctggg catggagagc aaggccatth ctgatgccc gatcactgcc 4500
 tccagctact tcaccaacat gtttgccacc tggagcccaa gcaaggccag gctgcacctc 4560
 cagggaagga gcaatgcctg gaggccccag gtcaacaacc caaaggagtg gctgcagggtg 4620
 gacttccaga agaccatgaa ggtcactggg gtgaccaccc aggggggtcaa gagcctgctc 4680
 accagcatgt atgtgaagga gttcctgac agctccagcc aggatggcca ccagtggacc 4740
 ctctttcttcc agaattggcaa ggtcaagggtg ttccagggca accaggacag cttcaccctt 4800
 gtgggtgaaca gcctggaccc cccctcctg accagatacc tgaggattca ccccagagc 4860
 tgggtccacc agattgccct gaggatggag gtcttgggat gtgaggccca ggacctgtac 4920
 tgatgaaata aaagatcttt attttcatta gatctgtgtg ttggtttttt gtgtgaggaa 4980

cccctagtga tggagttggc cactccctct ctgcgcgctc gctcgcctac tgaggccggg	5040
cgaccaaagg tgcgccgacg cccgggcttt gcccgggcgg cctcagttag cgagcgagcg	5100
cgcagagagg gaggggccaa gacgatttaa atgacaagct tggcgtaatc atggtcatag	5160
ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc	5220
ataaagtgtg aagcctgggg tgcctaataa gtgagctaac tcacattaat tgcgttgccg	5280
tcactgcccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa	5340
cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgtc cactgactcg	5400
ctgcgcctcg tegtccggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg	5460
ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag	5520
gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac	5580
gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga	5640
taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac cctgcgcgtt	5700
accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgcctttctca tagctcacgc	5760
tgtaggtatc tcagttcggg ttaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc	5820
ccggttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta	5880
agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat	5940
gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaagaaca	6000
gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct	6060
tgatccggca aacaaaccac cgctggtagc ggtgggtttt ttgtttgcaa gcagcagatt	6120
acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct	6180
cagtggaaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc	6240
acctagatcc ttttaaatta aaaatgaagt tttaaataca tctaaagtat atatgagtaa	6300
acttggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta	6360
tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc	6420
ttaccatctg gcccagtg cgcgaatgata ccgcgagacc cacgctcacc ggctccagat	6480
ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtgggtc tgcaacttta	6540
tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt	6600
aatagtttgc gcaacgttgt tgccattgct acaggcatcg tgggtgtcacg ctgcgtcgtt	6660

```

ggatatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg 6720
ttgtgcaaaa aagcgggttag ctccctcggt cctccgatac ttgtcagaag taagttggcc 6780
gcagtgttat cactcatggt tatggcagca ctgcataatt ctottactgt catgccatcc 6840
gtaagatgct tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg 6900
cggcgaccga gttgctcttg cccggcgctc atacgggata ataccgcgcc acatagcaga 6960
actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta 7020
ccgctgttga gatccagttc gatgtaacct actcgtgcac ccaactgata ttcagcatct 7080
tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag 7140
ggaataaggg cgacacggaa atgttgaata ctcatactct tcctttttca atattattga 7200
agcattttatc agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat 7260
aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc 7320
attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt tcgtctcgcg 7380
cgtttcggtg atgacggtga aaacctctga cacatgcagc tcccgggagac ggtcacagct 7440
tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc ggggtgttggc 7500
gggtgtcggg gctggcttaa ctatgcggca tcagagcaga ttgtactgag agtgcaccat 7560
atgcgggtgtg aaataccgca cagatgcgta aggagaaaat accgcatcag gcgccattcg 7620
ccattcaggc tgcgcaactg ttgggaaggg cgatcgggtc gggcctcttc gctattacgc 7680
cagctggcga aaggggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc 7740
cagtcacgac gttgtaaaac gacggccagt gaattcctcg agatttaaag gacg 7794

```

<210> 11
 <211> 15
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 11

Val	Ser	Asn	Asn	Val	Ser	Asn	Asn	Ala	Thr	Asn	Asn	Ala	Thr	Asn
1				5				10				15		

<210> 12
 <211> 45

<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 12
gtgagcaaca atgtgagcaa caatgccacc aataatgcta ccaac

45

<210> 13
<211> 11
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 13

Val Ser Asn Asn Ala Thr Asn Asn Val Ser Asn
1 5 10

<210> 14
<211> 33
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 14
gtgagcaaca atgccaccaa caatgtgagc aac

33

<210> 15
<211> 7
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 15

Val Ser Asn Asn Ala Thr Asn
1 5

<210> 16
<211> 21
<212> ADN
<213> Trình tự nhân tạo

<220>

<223> Tổng hợp

<400> 16

gtgagcaata atgccaccaa c

21

<210> 17

<211> 4

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Tổng hợp

<400> 17

Val Ser Asn Asn

1

<210> 18

<211> 12

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tổng hợp

<400> 18

gtgagcaata at

12

<210> 19

<211> 3

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Tổng hợp

<400> 19

Arg Ser Leu

1

<210> 20

<211> 9

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tổng hợp

<400> 20

aggagcctg

9

<210> 21
<211> 17
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 21

Ala Thr Asn Val Ser Asn Asn Ser Ala Thr Ser Ala Asp Ser Ala Val
1 5 10 15

Ser

<210> 22
<211> 51
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 22
gccactaatg tgtctaacaa ctctgctacc tctgctgact ctgctgtgag c

51

<210> 23
<211> 9
<212> PRT
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 23

Ala Thr Asn Tyr Val Asn Arg Ser Leu
1 5

<210> 24
<211> 27
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 24
gccaccaact atgtgaacag gagcctg

27

<210> 25
 <211> 21
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 25

Ala Thr Asn Tyr Val Asn Arg Ser Leu Ser Ala Thr Ser Ala Asp Ser
 1 5 10 15

Ala Val Ser Gln Asn
 20

<210> 26
 <211> 63
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 26

Gly Cys Cys Ala Cys Cys Ala Ala Cys Thr Ala Thr Gly Thr Gly Ala
 1 5 10 15

Ala Cys Ala Gly Gly Ala Gly Cys Cys Thr Gly Thr Cys Thr Gly Cys
 20 25 30

Cys Ala Cys Cys Thr Cys Thr Gly Cys Thr Gly Ala Cys Thr Cys Thr
 35 40 45

Gly Cys Thr Gly Thr Gly Ala Gly Cys Cys Ala Gly Ala Ala Thr
 50 55 60

<210> 27
 <211> 14
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 27

Val Ser Asn Asn Val Ser Asn Ala Val Ser Ala Val Ser Ala
 1 5 10

<210> 28
 <211> 42
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 28
 gtgagcaaca atgtgagcaa tgctgtgtct gctgtgtctg ct

42

<210> 29
 <211> 16
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 29

Ile Thr Val Ala Ser Ala Thr Ser Asn Ile Thr Val Ala Ser Ala Asp
 1 5 10 15

<210> 30
 <211> 48
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 30
 atcactgtgg cctctgccac ctctaacatc actgtggcct ctgctgac

48

<210> 31
 <211> 10
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 31

Ile Thr Val Thr Asn Ile Thr Val Thr Ala
 1 5 10

<210> 32
 <211> 30
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 32
 atcactgtga ccaacatcac tgtgactgcc

30

<210> 33
 <211> 10
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 33

Gln Thr Val Thr Asn Ile Thr Val Thr Ala
 1 5 10

<210> 34
 <211> 30
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 34
 cagactgtga ccaacatcac tgtgactgcc

30

<210> 35
 <211> 17
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 35

Ala Thr Asn Val Ser Asn Asn Ser Asn Thr Ser Asn Asp Ser Asn Val
 1 5 10 15

Ser

<210> 36
 <211> 51
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 36
 gccactaatg tgtctaacaa cagcaacacc agcaatgaca gcaatgtgtc t 51

<210> 37
 <211> 4374
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 37
 atgcagattg agctgagcac ctgcttcttc ctgtgcctgc tgaggttctg cttctctgcc 60
 accaggagat actacctggg ggctgtggag ctttcttggg actacatgca gtctgacctg 120
 ggggagctgc ctgtggatgc cagggtccca ccagagtgc ccaaatcctt cccattcaac 180
 acctctgtgg totacaagaa gaccctcttt gtggagttca ctgaccacct gttcaacatt 240
 gccaaaccca ggccaccctg gatgggactc ctgggacca ccattcaggc tgagggtgat 300
 gacactgtgg tcattaccct caagaacatg gcctcccacc ctgtgagcct gcatgctgtg 360
 ggggtcagct actggaaggc ctctgagggg gctgagtatg atgaccagac ctcccagagg 420
 gagaaggagg atgacaaagt gttccctggg ggcagccaca cctatgtgtg gcaggctctc 480
 aaggagaatg gcccctggc ctctgacca ctctgcctga cctactccta ctttctcat 540
 gtggacctgg tcaaggacct caactctgga ctgattgggg cctgctggt gtgcaggagg 600
 ggctccctgg ccaaagagaa gaccagacc ctgcacaagt tcattctcct gtttctgtgc 660
 tttgatgagg gcaagagctg gcaacttgaa accaagaact cctgatgca ggacagggat 720
 gctgcctctg ccagggctg gcccaagatg cacactgtga atggctatgt gaacaggagc 780
 ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcattgtgat tggcatgggg 840
 acaaccctg aggtgcactc cttttctctg gagggccaca ctttctggt caggaaccac 900
 agacaggcca gcctggagat cagcccatc accttctca ctgccagac cctgctgatg 960
 gacctcggac agttctctgt gttctgccac atcagctccc accagcatga tggcatggag 1020
 gcctatgtca aggtggacag ctgccctgag gagccacagc tcaggatgaa gaacaatgag 1080

gaggctgagg actatgatga tgacctgact gactctgaga tggatgtggt ccgctttgat	1140
gatgacaaca gcccatcctt cattcagatc aggtctgtgg ccaagaaaca cccaagacc	1200
tgggtgcact acattgctgc tgaggaggag gactgggact atgccccact ggtcctggcc	1260
cctgatgaca ggagctacaa gagccagtac ctcaacaatg gccacagag gattggacgc	1320
aagtacaaga aagtcagggt catggcctac actgatgaaa ccttcaagac caggagggcc	1380
attcagcatg agtctggcat cctgggcca ctctgtatg gggagggtgg ggacaccctg	1440
ctcatcatct tcaagaacca ggctccagg cctacaaca tctaccaca tggcatcact	1500
gatgtcaggc cctgtacag ccgcaggctg ccaaggggg tgaaacacct caaggacttc	1560
ccattctgc ctggggagat cttcaagtac aagtggactg tcaactgtga ggatggacca	1620
accaaactctg accccagggt cctcaccaga tactactoca gctttgtgaa catggagagg	1680
gacctggcct ctggcctgat tggccccactg ctcatctgct acaaggagtc tgtggaccag	1740
aggggaaacc agatcatgtc tgacaagagg aatgtgatc tgttctctgt ctttgatgag	1800
aacaggagct ggtacctgac tgagaacatt cagcgttcc tgcccaacc tgctggggtg	1860
cagctggagg acctgagtt ccaggccagc aacatcatgc actccatcaa tggctatgtg	1920
tttgacagcc tccagcttcc tgtctgctg catgagggtg cctactggta cattctttct	1980
attggggccc agactgactt cctttctgtc ttcttctctg gctacacctt caaacacaag	2040
atggtgtatg aggacacct gacctcttc ccattctctg gggagactgt gttcatgagc	2100
atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc	2160
atgactgcc tgctcaaagt ctctcctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacctgtc agcaagaaca atgccattga gccaggagc	2280
ttcagccaga atccacctgt cctgaaacgc caccagagg agatcaccag gaccacctc	2340
cagtctgacc aggaggagat tgactatgat gacaccattt ctgtggagat gaagaaagag	2400
gactttgaca tctatgacga ggacgagaac cagagcccaa ggagcttcca gaagaagacc	2460
aggcactact tcattgctgc tgtggagcgc ctgtgggact atggcatgag ctccagcccc	2520
catgtcctca ggaacagggc ccagtctggc tctgtgccac agttcaagaa agtggctctc	2580
caagagtcca ctgatggcag cttacccag cccctgtaca gaggggagct gaatgagcac	2640
ctgggactcc tgggcccata catcagggt gaggtggagg acaacatcat ggtgaccttc	2700
cgcaaccagg cctccaggcc ctacagcttc tacagctccc tcatcagcta tgaggaggac	2760

cagaggcagg gggctgagcc acgcaagaac ttgtgaaac ccaatgaaac caagacctac 2820
 ttctggaaaag tccagcacca catggccccc accaaggatg agtttgactg caaggcctgg 2880
 gctactttct ctgatgtgga cctggagaag gatgtgcaact ctggcctgat tggcccaactc 2940
 ctggtctgcc acaccaaac cctgaacct gcccatggaa ggcaagtgaac tgtgcaggag 3000
 tttgccctct tcttcacat ctttgatgaa accaagagct ggtacttcac tgagaacatg 3060
 gaggcgcaact gcaggccccc atgcaacatt cagatggagg accccacctt caaagagaac 3120
 taccgcttcc atgccatcaa tggctacatc atggacaccc tgcctgggct tgtcatggcc 3180
 caggaccaga ggatcagggtg gtacctgctt tctatgggct ccaatgagaa cattcactcc 3240
 atccacttct ctgggcatgt cttcactgtg cgcaagaagg aggagtaca gatggccctg 3300
 tacaacctct accctgggggt ctttgagact gtggagatgc tgcctccaa agctggcatc 3360
 tygaggggtg agtgccctcat tggggagcac ctgcatgctg gcatgagcac cctgttctctg 3420
 gtctacagca acaagtgcc aacccccctg ggaatggcct ctggccacat cagggacttc 3480
 cagatcactg cctctggcca gtatggccag tgggccccca agctggccag gctccactac 3540
 tctggatcca tcaatgcctg gagcaccaag gagccattca gctggatcaa agtggacctg 3600
 ctggccccca tgatcatcca tggcatcaag acccaggggg ccaggcagaa gttctccagc 3660
 ctgtacatca gccagttcat catcatgtac agcctggatg gcaagaaatg gcagacctac 3720
 agaggcaact ccactggaac actcatggtc ttctttggca atgtggacag ctctggcatc 3780
 aagcacaaca tcttcaacct cccaatcctc gccagatata tcaggctgca cccacccac 3840
 tacagcatcc gcagcacct caggatggag ctgatgggct gtgacctgaa ctctgcagc 3900
 atgcccctgg gcatggagag caaggccatt tctgatgcc agatcactgc ctccagctac 3960
 ttcaccaaca tgtttgccac ctggagccca agcaaggcca ggctgcacct ccagggaagg 4020
 agcaatgcct ggaggcccca ggtcaacaac ccaaaggagt ggctgcagggt ggacttccag 4080
 aagaccatga aggtcactgg ggtgaccacc caggggggtca agagcctgct caccagcatg 4140
 tatgtgaagg agttctgat cagctccagc caggatggcc accagtggac cctcttcttc 4200
 cagaatggca aggtcaagggt gttccagggc aaccaggaca gcttcacccc tgtggtgaac 4260
 agcctggacc cccccctct gaccagatag ctgaggattc accccagag ctgggtccac 4320
 cagattgccc tgaggatgga ggtcctggga tgtgaggccc aggacctgta ctga 4374

<210> 38
 <211> 5191
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 38
 ttggccactc cctctctgcg cgctcgctcg ctcaactgagg ccgggcgacc aaaggctcgcc 60
 cgaagccccg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agaggggagtg 120
 gccaaactcca tcaactagggg ttctctgagtt taaacttcgt cgacggggga ggctgctggg 180
 gaatattaac caaggtcacc ccagttatcg gaggagcaaa caggggctaa gtccaccggg 240
 ggaggtctgt ggtgaatatt aaccaaggtc accccagtta tcggaggagc aaacaggggc 300
 taagtccaca aatgacctat taagaatatt tcatagaacg aatgttccga tgctctaato 360
 totctagaca aggttcatat ttgtatgggt tacttatctt ctctttgttg actaagtcaa 420
 taatcagaat cagcagggtt gcagtcagat tggcagggat aagcagccta gctcaggaga 480
 agtgagtata aaagccccag gctgggagca gccatcacag aagtccactc attcttggca 540
 ggggcgcgcc gccaccatgc agattgagct gagcacctgc ttcttctgtt gcctgctgag 600
 gttctgcttc tctgccacca ggagatacta cctgggggct gtggagcttt cttgggacta 660
 catgcagtct gacctggggg agctgcctgt ggatgccagg tccccacca gagtgcccaa 720
 atccttccca ttcaaacact ctgtggtcta caagaagacc ctctttgttg agttcactga 780
 ccacctgttc aacattgcc aaccaggcc accctggatg ggactcctgg gaccaccat 840
 tcaggctgag gtgtatgaca ctgtggtcgt caccctcaag aacatggcct cccaccctgt 900
 gagcctgcat gctgtggggg tcagctactg gaagtcctct gagggggctg agtatgatga 960
 ccagacctcc cagagggaga aggaggatga caaagtgttc cctgggaaga gccacaccta 1020
 tgtgtggcag gtccctcaagg agaatggccc cactgcctct gaccaccctt gcctgaccta 1080
 ctctacctt tctcatgttg acctgggtcaa ggacctcaac tctggactga ttggggccct 1140
 gctggtgtgc agggagggct cctgggcaa agagaagacc cagaccctgc acaagttcat 1200
 tctctgtttt gctgtctttg atgagggcaa gagctggcac tctgaaacca agaactccct 1260
 gatgcaggac agggatgctg cctctgccag ggctggccc aagatgcaca ctgtgaatgg 1320
 ctatgtgaac aggagcctgc ctggactcat tggctgccac aggaaatctg tctactggca 1380
 tgtgattggc atgggggaaa cccctgaggt gcactccatt ttctggagg gccacacctt 1440

cctggtcagg aaccacagac aggccagcct ggagatcagc cccatcacct tcctcactgc	1500
ccagaccctg ctgatggacc tcggacagtt cctgctgttc tgccacatca gctcccacca	1560
gcatgatggc atggaggcct atgtcaaggt ggacagctgc cctgaggagc cacagctcag	1620
gatgaagaac aatgaggagg ctgaggacta tgatgatgac ctgactgact ctgagatgga	1680
tgtgggtccgc tttgatgatg acaacagccc atccttcatt cagatcaggt ctgtggccaa	1740
gaaacacccc aagacctggg tgcactacat tgctgctgag gaggaggact gggactatgc	1800
cccactggtc ctggcccctg atgacaggag ctacaagagc cagtacctca acaatggccc	1860
acagaggatt ggacgcaagt acaagaaagt caggttcatg gcctacactg atgaaacctt	1920
caagaccagg gaggccattc agcatgagtc tggcatcctg ggcccactcc tgtatgggga	1980
ggtggggggac accctgctca tcattctcaa gaaccaggcc tccaggccct acaacatcta	2040
cccacatggc atcactgatg tcaggcccct gtacagccgc aggctgccaa agggggtgaa	2100
acacctcaag gacttcccca ttctgcctgg ggagatcttc aagtacaagt ggactgtcac	2160
tgtggaggat ggaccaacca aatctgacct cagggtgcctc accagatact actccagctt	2220
tgtgaacatg gagagggacc tggcctctgg cctgattggc cactgctca tctgctacaa	2280
ggagtctgtg gaccagaggg gaaaccagat catgtctgac aagaggaatg tgattctgtt	2340
ctctgtcttt gatgagaaca ggagctggta cctgactgag aacattcagc gcttccctgcc	2400
caaccctgct ggggtgcagc tggaggacct tgagttccag gccagcaaca tcatgcactc	2460
catcaatggc tatgtgtttg acagcctcca gttttctgtc tgcttgcctg aggtggccta	2520
ctggtacatt ctttctattg gggcccagac tgacttccct tctgtcttct tctctggcta	2580
caccttcaaa cacaagatgg tgtatgagga caccctgacc ctcttcccat tctctgggga	2640
gactgtgttc atgagcatgg agaaccctgg cctgtggatt ctgggatgcc acaactctga	2700
cttccgcaac aggggcatga ctgccctgct caaagtctcc tcctgtgaca agaactctgg	2760
ggactactat gaggacagct atgaggacat ctctgcctac ctgctcagca agaacaatgc	2820
cattgagccc aggagcttca gccagaatgt gagcaataat gccaccaacc cacctgtcct	2880
gaaacgccac cagagggaga tcaccaggac caccctccag tctgaccagg aggagattga	2940
ctatgatgac accatttctg tggagatgaa gaaagaggac tttgacatct atgacgagga	3000
cgagaaccag agcccaagga gcttccagaa gaagaccagg cactacttca ttgctgctgt	3060
ggagcgccctg tgggactatg gcatgagctc cagcccccat gtccctcagga acagggccca	3120

gtctggctct gtgccacagt tcaagaaagt ggtcttccaa gagttcactg atggcagctt 3180
 caccagccc ctgtacagag gggagctgaa tgagcacctg ggactcctgg gccatacat 3240
 cagggtgag gtggaggaca acatcatggt gaccttccgc aaccaggcct ccaggcccta 3300
 cagcttctac agctccctca tcagctatga ggaggaccag aggcaggggg ctgagccacg 3360
 caagaacttt gtgaaaccca atgaaaccaa gacctacttc tggaaagtcc agcaccacat 3420
 ggccccacc aaggatgagt ttgactgcaa ggctggggc tacttctctg atgtggacct 3480
 ggagaaggat gtgcactctg gcctgattgg cccactcctg gtctgccaca ccaacaccct 3540
 gaacctgcc catggaagge aagtgactgt gcaggagttt gccctcttct tcaccatctt 3600
 tgatgaaacc aagagctggt acttcaactga gaacatggag cgcaactgca gggccccatg 3660
 caacattcag atggaggacc ccaccttcaa agagaactac cgcttccatg ccatcaatgg 3720
 ctacatcatg gacaccctgc ctgggcttgt catggcccag gaccagagga tcagggtgta 3780
 cctgctttct atgggctcca atgagaacat tcactccatc cacttctctg ggcatgtctt 3840
 cactgtgcgc aagaaggagg agtacaagat ggccctgtac aacctctacc ctggggctct 3900
 tgagactgtg gagatgtgc cctccaaagc tggcatctgg aggggtggagt gcctcattgg 3960
 ggagcacctg catgctggca tgagcacctt gttcctggtc tacagcaaca agtgccagac 4020
 cccctggga atggcctctg gccacatcag ggacttccag atcactgcct ctggccagta 4080
 tggccagtgg gccccaaagc tggccaggct ccactactct ggatccatca atgcctggag 4140
 caccaaggag ccattcagct ggatcaaagt ggacctgctg gccccatga tcatccatgg 4200
 catcaagacc cagggggcca ggcagaagtt ctccagcctg tacatcagcc agttcatcat 4260
 catgtacagc ctggatggca agaaatggca gacctacaga ggcaactcca ctggaacact 4320
 catggtcttc ttggaatg tggacagctc tggcatcaag cacaacatct tcaaccccc 4380
 aatcatcgcc agatacatca ggctgcacc caccactac agcatccgca gcacctcag 4440
 gatggagctg atgggctgtg acctgaactc ctgcagcatg ccctgggca tggagagcaa 4500
 ggccatttct gatgccaga tcactgcctc cagctacttc accaacadgt ttgccacctg 4560
 gagccaagc aaggccaggc tgcacctcca gggaaggagc aatgcctgga ggccccaggt 4620
 caacaaccca aaggagtggc tgcaggtgga cttccagaag accatgaagg tcactggggt 4680
 gaccaccag ggggtcaaga gcctgctcac cagcatgtat gtgaaggagt tcctgatcag 4740
 ctccagccag gatggccacc agtggacctt cttcttccag aatggcaagg tcaaggtgtt 4800

ccagggcaac caggacagct tcacccctgt ggtgaacagc ctggaccccc cctcctgac 4860
 cagatacctg aggattcacc ccagagagctg ggtccaccag attgccctga ggatggaggt 4920
 cctgggatgt gaggcccagg acctgtactg atgagcggcc gctcttagta gcagtatcga 4980
 taataaaaga tctttatctt cattagatct gtgtgttggg tttttgtgtg ttaattaagc 5040
 tcgcgaagga acccctagtg atggagttag ccactccctc tctgcgcgct cgctcgctca 5100
 ctgagggccgg gcgaccaaag gtcgcccagc gcccgggctt tgcccgggcg gcctcagtga 5160
 gcgagcgagc gcgcagagag ggagtggcca a 5191

<210> 39
 <211> 2351
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 39

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
 1 5 10 15

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

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
 35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
 50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
 65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
 85 90 95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser
 100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
 115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
165 170 175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
180 185 190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
195 200 205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
305 310 315 320

Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
325 330 335

Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
340 345 350

Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365

Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380

Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400

Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415

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

Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
 435 440 445

Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu
 450 455 460

Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
 465 470 475 480

Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
 485 490 495

His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys
 500 505 510

Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe
 515 520 525

Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp
 530 535 540

Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg
 545 550 555 560

Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu
 565 570 575

Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val
580 585 590

Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu
595 600 605

Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp
610 615 620

Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val
625 630 635 640

Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp
645 650 655

Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe
660 665 670

Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr
675 680 685

Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
690 695 700

Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
705 710 715 720

Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
725 730 735

Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys
740 745 750

Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro
755 760 765

Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp
770 775 780

Ile Glu Lys Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys
785 790 795 800

Ile Gln Asn Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser
805 810 815

Pro Thr Pro His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr
820 825 830

Glu Thr Phe Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn
835 840 845

Ser Leu Ser Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly
850 855 860

Asp Met Val Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu
865 870 875 880

Lys Leu Gly Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys
885 890 895

Val Ser Ser Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn
900 905 910

Leu Ala Ala Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met
915 920 925

Pro Val His Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys
930 935 940

Ser Ser Pro Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu
945 950 955 960

Asn Asn Asp Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu
965 970 975

Ser Ser Trp Gly Lys Asn Val Ser Ser Thr Glu Ser Gly Arg Leu Phe
980 985 990

Lys Gly Lys Arg Ala His Gly Pro Ala Leu Leu Thr Lys Asp Asn Ala
995 1000 1005

Leu Phe Lys Val Ser Ile Ser Leu Leu Lys Thr Asn Lys Thr Ser
1010 1015 1020

Asn Asn Ser Ala Thr Asn Arg Lys Thr His Ile Asp Gly Pro Ser
 1025 1030 1035
 Leu Leu Ile Glu Asn Ser Pro Ser Val Trp Gln Asn Ile Leu Glu
 1040 1045 1050
 Ser Asp Thr Glu Phe Lys Lys Val Thr Pro Leu Ile His Asp Arg
 1055 1060 1065
 Met Leu Met Asp Lys Asn Ala Thr Ala Leu Arg Leu Asn His Met
 1070 1075 1080
 Ser Asn Lys Thr Thr Ser Ser Lys Asn Met Glu Met Val Gln Gln
 1085 1090 1095
 Lys Lys Glu Gly Pro Ile Pro Pro Asp Ala Gln Asn Pro Asp Met
 1100 1105 1110
 Ser Phe Phe Lys Met Leu Phe Leu Pro Glu Ser Ala Arg Trp Ile
 1115 1120 1125
 Gln Arg Thr His Gly Lys Asn Ser Leu Asn Ser Gly Gln Gly Pro
 1130 1135 1140
 Ser Pro Lys Gln Leu Val Ser Leu Gly Pro Glu Lys Ser Val Glu
 1145 1150 1155
 Gly Gln Asn Phe Leu Ser Glu Lys Asn Lys Val Val Val Gly Lys
 1160 1165 1170
 Gly Glu Phe Thr Lys Asp Val Gly Leu Lys Glu Met Val Phe Pro
 1175 1180 1185
 Ser Ser Arg Asn Leu Phe Leu Thr Asn Leu Asp Asn Leu His Glu
 1190 1195 1200
 Asn Asn Thr His Asn Gln Glu Lys Lys Ile Gln Glu Glu Ile Glu
 1205 1210 1215
 Lys Lys Glu Thr Leu Ile Gln Glu Asn Val Val Leu Pro Gln Ile
 1220 1225 1230

His Thr Val Thr Gly Thr Lys Asn Phe Met Lys Asn Leu Phe Leu
 1235 1240 1245

Leu Ser Thr Arg Gln Asn Val Glu Gly Ser Tyr Asp Gly Ala Tyr
 1250 1255 1260

Ala Pro Val Leu Gln Asp Phe Arg Ser Leu Asn Asp Ser Thr Asn
 1265 1270 1275

Arg Thr Lys Lys His Thr Ala His Phe Ser Lys Lys Gly Glu Glu
 1280 1285 1290

Glu Asn Leu Glu Gly Leu Gly Asn Gln Thr Lys Gln Ile Val Glu
 1295 1300 1305

Lys Tyr Ala Cys Thr Thr Arg Ile Ser Pro Asn Thr Ser Gln Gln
 1310 1315 1320

Asn Phe Val Thr Gln Arg Ser Lys Arg Ala Leu Lys Gln Phe Arg
 1325 1330 1335

Leu Pro Leu Glu Glu Thr Glu Leu Glu Lys Arg Ile Ile Val Asp
 1340 1345 1350

Asp Thr Ser Thr Gln Trp Ser Lys Asn Met Lys His Leu Thr Pro
 1355 1360 1365

Ser Thr Leu Thr Gln Ile Asp Tyr Asn Glu Lys Glu Lys Gly Ala
 1370 1375 1380

Ile Thr Gln Ser Pro Leu Ser Asp Cys Leu Thr Arg Ser His Ser
 1385 1390 1395

Ile Pro Gln Ala Asn Arg Ser Pro Leu Pro Ile Ala Lys Val Ser
 1400 1405 1410

Ser Phe Pro Ser Ile Arg Pro Ile Tyr Leu Thr Arg Val Leu Phe
 1415 1420 1425

Gln Asp Asn Ser Ser His Leu Pro Ala Ala Ser Tyr Arg Lys Lys
 1430 1435 1440

Asp Ser Gly Val Gln Glu Ser Ser His Phe Leu Gln Gly Ala Lys
1445 1450 1455

Lys Asn Asn Leu Ser Leu Ala Ile Leu Thr Leu Glu Met Thr Gly
1460 1465 1470

Asp Gln Arg Glu Val Gly Ser Leu Gly Thr Ser Ala Thr Asn Ser
1475 1480 1485

Val Thr Tyr Lys Lys Val Glu Asn Thr Val Leu Pro Lys Pro Asp
1490 1495 1500

Leu Pro Lys Thr Ser Gly Lys Val Glu Leu Leu Pro Lys Val His
1505 1510 1515

Ile Tyr Gln Lys Asp Leu Phe Pro Thr Glu Thr Ser Asn Gly Ser
1520 1525 1530

Pro Gly His Leu Asp Leu Val Glu Gly Ser Leu Leu Gln Gly Thr
1535 1540 1545

Glu Gly Ala Ile Lys Trp Asn Glu Ala Asn Arg Pro Gly Lys Val
1550 1555 1560

Pro Phe Leu Arg Val Ala Thr Glu Ser Ser Ala Lys Thr Pro Ser
1565 1570 1575

Lys Leu Leu Asp Pro Leu Ala Trp Asp Asn His Tyr Gly Thr Gln
1580 1585 1590

Ile Pro Lys Glu Glu Trp Lys Ser Gln Glu Lys Ser Pro Glu Lys
1595 1600 1605

Thr Ala Phe Lys Lys Lys Asp Thr Ile Leu Ser Leu Asn Ala Cys
1610 1615 1620

Glu Ser Asn His Ala Ile Ala Ala Ile Asn Glu Gly Gln Asn Lys
1625 1630 1635

Pro Glu Ile Glu Val Thr Trp Ala Lys Gln Gly Arg Thr Glu Arg
1640 1645 1650

Leu Cys Ser Gln Asn Pro Pro Val Leu Lys Arg His Gln Arg Glu
 1655 1660 1665

Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr
 1670 1675 1680

Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile
 1685 1690 1695

Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys
 1700 1705 1710

Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr
 1715 1720 1725

Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser
 1730 1735 1740

Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr
 1745 1750 1755

Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu
 1760 1765 1770

His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp
 1775 1780 1785

Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser
 1790 1795 1800

Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly
 1805 1810 1815

Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr
 1820 1825 1830

Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu
 1835 1840 1845

Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu
 1850 1855 1860

Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His
1865 1870 1875

Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln
1880 1885 1890

Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp
1895 1900 1905

Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn
1910 1915 1920

Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His
1925 1930 1935

Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met
1940 1945 1950

Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser
1955 1960 1965

Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr
1970 1975 1980

Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr
1985 1990 1995

Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly
2000 2005 2010

Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly
2015 2020 2025

Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro
2030 2035 2040

Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala
2045 2050 2055

Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His
2060 2065 2070

Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser
2075 2080 2085

Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile
2090 2095 2100

Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser
2105 2110 2115

Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr
2120 2125 2130

Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn
2135 2140 2145

Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile
2150 2155 2160

Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg
2165 2170 2175

Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys
2180 2185 2190

Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln
2195 2200 2205

Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser
2210 2215 2220

Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp
2225 2230 2235

Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe
2240 2245 2250

Gln Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys
2255 2260 2265

Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser
2270 2275 2280

Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys
2285 2290 2295

Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val
2300 2305 2310

Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile His
2315 2320 2325

Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met Glu Val Leu
2330 2335 2340

Gly Cys Glu Ala Gln Asp Leu Tyr
2345 2350

<210> 40
<211> 42
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 40
agcttcagcc agaatccacc tgtcctgaaa cgccaccaga gg 42

<210> 41
<211> 87
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<400> 41
agcttcagcc agaatgtgag caacaatgtg agcaacaatg ccaccaataa tgctaccaac 60

ccacctgtcc tgaaacgcc caagagg 87

<210> 42
<211> 75
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Tổng hợp

<210> 47
 <211> 69
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 47
 agcttcagcc agaatgccac caactatgtg aacaggagcc tgccacctgt cctgaaacgc 60
 caccagagg 69

<210> 48
 <211> 105
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 48
 agcttcagcc agaatgccac caactatgtg aacaggagcc tgtctgccac ctctgtgtgac 60
 tctgtgtgtga gccagaatcc acctgtcctg aaacgccacc agagg 105

<210> 49
 <211> 84
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 49
 agcttcagcc agaatgtgag caacaatgtg agcaatgctg tgtctgtgtgt gtctgtctca 60
 cctgtcctga aacgccacca gagg 84

<210> 50
 <211> 90
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 50
 agcttcagcc agaatatcac tgtggcctct gccacctcta acatcactgt ggctctgtgt 60
 gaccacactg tctgaaacg ccaccagagg 90

<400> 42
 agcttcagcc agaattgtgag caacaatgcc accaacaatg tgagcaacc acctgtctg 60
 aaacgccacc agagg 75

<210> 43
 <211> 63
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 43
 agcttcagcc agaattgtgag caataatgcc accaaccac ctgtcctgaa acgccaccag 60
 agg 63

<210> 44
 <211> 54
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 44
 agcttcagcc agaattgtgag caataatcca cctgtcctga aacgccacca gagg 54

<210> 45
 <211> 51
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 45
 agcttcagcc agaataggag cctgccacct gtctgaaac gccaccagag g 51

<210> 46
 <211> 93
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 46
 agcttcagcc agaatgccac taatgtgtct aacaactctg ctacctctgc tgactctgct 60
 gtgagccac ctgtcctgaa acgccaccag agg 93

<210> 51
 <211> 72
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 51
 agcttcagcc agaatatcac tgtgaccaac atcactgtga ctgccccacc tgtcctgaaa 60
 cgccaccaga gg 72

<210> 52
 <211> 72
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 52
 agcttcagcc agaatcagac tgtgaccaac atcactgtga ctgccccacc tgtcctgaaa 60
 cgccaccaga gg 72

<210> 53
 <211> 93
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 53
 agcttcagcc agaatgccac taatgtgtct aacaacagca acaccagcaa tgacagcaat 60
 gtgtctccac ctgtcctgaa acgccaccag agg 93

<210> 54
 <211> 2674
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 54
 gacgatttaa atgacaagct tggcgtaatc atgggtcatag ctgtttcctg tgtgaaattg 60
 ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgtg aagcctgggg 120

tgctaata ga gtgagcta ac tcacatta at tgcgttgccg tcaactgccc ctttccagtc	180
gggaaacctg tcgtgccagc tgcattaatg aatcgcccaa cgcgcgggga gaggcggttt	240
gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct	300
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga	360
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc	420
cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg	480
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt tccccctgg	540
aagctccctc gtgcgctctc ctgttccgac cctgcgcgtt accggatacc tgtccgcctt	600
tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggc	660
gtaggtcgtt cgctccaagc tgggctgtgt gcacgaacct cccgttcagc ccgaccgctg	720
cgccttatcc ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact	780
ggcagcagcc actggtaaca ggattagcag agcgaggatg gtaggcggtg ctacagagtt	840
cttgaagtgg tggcctaact acggctacac tagaagaaca gtatttggtg tctgcgctct	900
gotgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac	960
cgcctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc	1020
tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg	1080
ttaagggatt ttggatcatg gattatcaaa aaggatcttc acctagatcc ttttaaatta	1140
aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggctctg acagttacca	1200
atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc	1260
ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc	1320
tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc	1380
agccggaagg gccgagcgca gaagtggtec tgcaacttta tccgcctcca tccagtctat	1440
taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacggtgt	1500
tgccattgct acaggcatcg tgggtgcacg ctcgctgctt ggtatggctt cattcagctc	1560
cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag	1620
ctccttcggc cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt	1680
tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac	1740
tggtagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg	1800

ccggggtca atacgggata ataccggccc acatagcaga actttaaaag tgctcatcat 1860
 tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc 1920
 gatgtaaccc actcgtgcac ccaactgata ttcagcatct tttactttca ccagcgtttc 1980
 tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa 2040
 atgttgaata ctcatactct tcttttttca atattattga agcatttatc agggttattg 2100
 tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg 2160
 cacatttccc cgaaaagtgc cacctgacgt ctaagaaacg attattatca tgacattaac 2220
 ctataaaaaat aggcgtatca cgaggccctt tcgtctcgcg cgtttcggtg atgacggtga 2280
 aaacctctga cacatgcagc tcccggagac ggtcacagct tgtctgtaag cggatgcgg 2340
 gagcagacaa gcccgtcagg gcgcgtcagc ggggtgttggc ggggtgtcggg gctggcttaa 2400
 ctatgcggca tcagagcaga ttgtactgag agtgcaccat atgcggtgtg aaataccgca 2460
 cagatgcgta aggagaaaaat accgcatcag gcgccattcg ccattcaggc tgcgcaactg 2520
 ttgggaaggg cgatcgggtgc gggcctcttc gtattacgc cagctggcga aagggggatg 2580
 tgctgcaagg cgattaagtt gggtaacgcc agggtttttc cagtcacgac gttgtaaaac 2640
 gacggccagt gaattcctcg agatttaaatt gacg 2674

<210> 55

<211> 1015

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tổng hợp

<400> 55

tggcgtaatc atggcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac 60
 acaacatacg agccggaagc ataaagtgtg aagcctgggg tgccaatga gtgagctaac 120
 tcacattaat tgcgttgccg tcaactgccg ctttccagtc gggaaacctg tcgtgccagc 180
 tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgtctttccg 240
 cttctcgcgt cactgactcg ctgcgctcgg tcgttcgggt gcggcgagcg gtatcagctc 300
 actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 360
 gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 420
 ataggtcccg ccccccagc gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 480

acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc 540
 ctgttccgac cctgcgcgtt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 600
 cgctttctca tagctcacgc tgtaggtatc tcagttcggg gtaggtcgtt cgctccaagc 660
 tgggctgtgt gcacgaaccc ccggttcagc ccgacgcgtg cgccttatcc ggtaactatc 720
 gtcttgagtc caaccggta agacacgact tategccact ggcagcagcc actggtaaca 780
 ggattagcag agcgaggtat gtaggcgggtg ctacagagtt ctggaagtgg tggcctaact 840
 accgctacac tagaagaaca gtatttggtg tctgcgctct gotgaagcca gttaccttcg 900
 gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtgggtttt 960
 ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat ccttt 1015

<210> 56
 <211> 861
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Tổng hợp

<400> 56
 ttaccaatgc ttaatcagtg aggcacctat ctacgcgac tgtctatttc gttcatccat 60
 agttgcctga ctccccgtcg ttagataaac tacgatacgg gagggcttac catctggccc 120
 cagtgtgca atgataccgc gagaccacg ctgacgggct ccagatttat cagcaataaa 180
 ccagccagcc ggaagggccg agcgcagaag tggctctgca actttatccg cctccatcca 240
 gtctattaat tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa 300
 cgttgttgcc attgctacag gcacgtggg gtcacgctcg tcgtttggta tgggttcatt 360
 cagctccggg tcccaacgat caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc 420
 ggttagctcc ttcggctctc cgatcgttgt cagaagtaag ttggccgcag tgttatcact 480
 catggttatg gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgcttttc 540
 tgtgactggg gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg 600
 ctcttgcccg gcgtcaatac gggataatac cgcgccacat agcagaaact taaaagtgt 660
 catcattgga aaacgttctt cggggcgaaa actctcaagg atcttaccgc tgttgagatc 720
 cagttcgatg taaccactc gtgcacccaa ctgatcttca gcactttta ctttcaccag 780
 cgtttctggg tgagcaaaaa caggaaggca aaatgccgca aaaaagggaa taagggcgac 840

acggaaatgt tgaataactca t

861

<210> 57

<211> 7865

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Tổng hợp

<400> 57

tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat	300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cctcgagatt taaatgacgt	420
tggccaactcc ctctctgcgc gctcgtcgc tctactgaggc cgggcgacca aaggtcgccc	480
gacgcccggg ctttgcccg gcggcctcag tgagcgagcg agcgcgaga gagggagtgg	540
ccaactccat cactaggggt tcttgagttt aaacttcgtc gacgggggag gctgctggtg	600
aatattaacc aaggtcaccc cagttatcgg aggagcaaac aggggctaag tccaccgggg	660
gaggctgctg gtgaatatta accaagggtca cccagttat cggaggagca aacaggggct	720
aagtccacaa atgacctatt aagaatat tcatagaacga atgttccgat gctctaattct	780
ctctagacaa gggtcatatt tgtatgggtt acttattctc tctttgttga ctaagtcaat	840
aatcagaatc agcaggtttg cagtcagatt ggcagggata agcagcctag ctcaggagaa	900
gtgagtataa aagccccagg ctgggagcag ccatcacaga agtccactca ttcttggcag	960
gggcgcgccg ccaccatgca gattgagctg agcacctgct tcttcctgtg cctgctgagg	1020
ttctgcttct ctgccaccag gagatactac ctgggggctg tggagcttct ttgggactac	1080
atgcagtctg acctggggga gctgcctgtg gatgccaggt tcccaccag agtgcccaaa	1140
tccttcccat tcaacacctc tgtggtctac aagaagacco tctttgtgga gttcactgac	1200
cacctgttca acattgccaa acccaggcca cctggatgg gactcctggg acccaccatt	1260
caggctgagg tgtatgacac tgtggtcgtc accctcaaga acatggcctc ccacctgtg	1320

agcctgcatg ctgtgggggt cagctactgg aagtcctctg agggggctga gtatgatgac 1380
 cagacctccc agaggagaa ggaggatgac aaagtgttcc ctgggaagag ccacacctat 1440
 gtgtggcagg tcctcaagga gaatggcccc actgcctctg acccacctg cctgacctac 1500
 tcctaccttt ctcatgtgga cctgggtcaag gacctcaact ctggactgat tggggccctg 1560
 ctgggtgtgca gggagggctc cctggccaaa gagaagaccc agacctgca caagttcatt 1620
 ctctgttttg ctgtctttga tgagggcaag agctggcact ctgaaaccaa gaactccctg 1680
 atgcaggaca gggatgctgc ctctgccagg gcctggccca agatgcacac tgtgaatggc 1740
 tatgtgaaca ggagcctgcc tggactcatt ggctgccaca ggaaatctgt ctactggcat 1800
 gtgattggca tggggacaac cctgaggtg cactccattt tcctggaggg ccacaccttc 1860
 ctggtcagga accacagaca ggccagcctg gagatcagcc ccacacctt cctcactgcc 1920
 cagacctgc tgatggacct cggacagttc ctgctgttct gccacatcag cccccaccag 1980
 catgatggca tggaggccta tgtcaaggtg gacagctgcc ctgaggagcc acagctcagg 2040
 atgaagaaca atgaggaggc tgaggactat gatgatgacc tgactgactc tgagatggat 2100
 gtggtccgct ttgatgatga caacagccca tcttcattc agatcaggtc tgtggccaag 2160
 aaacacccca agacctgggt gcactacatt gctgctgagg aggaggactg ggactatgcc 2220
 ccactggctc tggccctga tgacaggagc tacaagagcc agtacctcaa caatggccca 2280
 cagaggattg gacgcaagta caagaaagtc aggttcatgg cctacactga tgaaaccttc 2340
 aagaccaggg aggccattca gcatgagtct ggcacctgg gccactcct gtatggggag 2400
 gtgggggaca ccctgctcat catcttcaag aaccaggcct ccaggcccta caacatctac 2460
 ccacatggca tcactgatgt caggccctg tacagccgca ggctgccaaa gggggtgaaa 2520
 cacctcaagg acttccccat tctgcctggg gagatcttca agtacaagtg gactgtcact 2580
 gtggaggatg gaccaaccaa atctgacccc aggtgcctca ccagatacta ctccagcttt 2640
 gtgaacatgg agagggacct ggctctggc ctgattggcc cactgctcat ctgctacaag 2700
 gagtctgtgg accagagggg aaaccagatc atgtctgaca agaggaatgt gattctgttc 2760
 tctgtctttg atgagaacag gagctggtac ctgactgaga acattcagcg cttcctgccc 2820
 aacctgctg ggggtcagct ggaggacct gagttccagg ccagcaacat catgcactcc 2880
 atcaatggct atgtgtttga cagcctccag ctttctgtct gcctgcatga ggtggcctac 2940
 tggtaattc tttctattgg ggcccagact gacttccttt ctgtcttctt ctctggctac 3000

accttcaaac acaagatggt gtatgaggac accctgaacc ttttccatt ctctggggag	3060
actgtgttca tgagcatgga gaaccctggc ctgtggattc tgggatgcca caactctgac	3120
ttccgcaaca ggggcatgac tgccctgctc aaagtctcct cctgtgacaa gaacactggg	3180
gactactatg aggacagcta tgaggacatc totgcctacc tgctcagcaa gaacaatgcc	3240
attgagccca ggagcttcag ccagaatgtg agcaataatg ccaccaaccc acctgtcctg	3300
aaacgccacc agagggagat caccaggacc accctccagt ctgaccagga ggagattgac	3360
tatgatgaca ccatttctgt ggagatgaag aaagaggact ttgacatcta tgacgaggac	3420
gagaaccaga gccaaggag ctccagaag aagaccaggc actacttcat tgctgctgtg	3480
gagcgctgtt gggactatgg catgagctcc agccccatg tcttcaggaa cagggcccag	3540
tctggctctg tgccacagtt caagaaagtg gtcttccaag agttcactga tggcagcttc	3600
accagcccc tgtacagagg ggagctgaat gagcacctgg gactcctggg ccatacatc	3660
agggtgagg tggaggacaa catcatggtg acctccgca accaggcctc caggccctac	3720
agcttctaca gctccctcat cagctatgag gaggaccaga ggcagggggc tgagccacgc	3780
aagaactttg tgaaacccaa tgaaaccaag acctacttct ggaaagtcca gcaccacatg	3840
gccccacca aggatgagtt tgactgcaag gcttgggctt acttctctga tgtggacctg	3900
gagaaggatg tgcactctgg cctgattggc ccactcctgg tctgccacac caacacctg	3960
aacctgccc atggaaggca agtgactgtg caggagtttg cctcttctt caccatcttt	4020
gatgaaacca agagctggtg ctctactgag aacatggagc gcaactgcag ggccccatgc	4080
aacattcaga tggaggaccc caccttcaaa gagaactacc gcttccatgc catcaatggc	4140
tacatcatgg acacctgcc tgggcttgct atggcccagg accagaggat cagggtggtac	4200
ctgctttcta tgggctccaa tgagaacatt cactccatcc acttctctgg gcatgtcttc	4260
actgtgcgca agaaggagga gtacaagatg gccctgtaca acctctaccc tggggtcttt	4320
gagactgtgg agatgctgcc ctccaaagct ggcatctgga ggggtggagt cctcattggg	4380
gagcacctgc atgctggcat gagcacctg ttcttgggtc acagcaacaa gtgccagacc	4440
ccctgggaa tggcctctgg ccacatcagg gacttcaga tcaactgcctc tggccagtat	4500
ggccagtggg cccccaagct ggccaggctc cactactctg gatccatcaa tgctggagc	4560
accaaggagc cattcagctg gatcaaagtg gacctgctgg ccccatgat catccatggc	4620
atcaagaccc agggggccag gcagaagttc tccagcctgt acatcagcca gttcatcatc	4680

atgtacagcc tggatggcaa gaaatggcag acctacagag gcaactccac tggaaacactc	4740
atggtcttct tggcaatgt ggacagctct ggcaccaage acaacatctt caacccccca	4800
atcatcgcca gatacatcag gctgcacccc acccaactaca gcatccgcag caccctcagg	4860
atggagctga tgggctgtga cctgaactcc tgcagcatgc ccctgggcat ggagagcaag	4920
gccatttctg atgcccagat cactgcctcc agctacttca ccaacatggt tgcacactgg	4980
agcccaagca aggccaggct gcacctccag ggaaggagca atgcctggag gccccaggtc	5040
aacaacccaa aggagtggct gcagggtggac ttccagaaga ccatgaaggc cactgggggtg	5100
accacccagg gggcacaagag cctgctcacc agcatgtatg tgaaggagtt cctgatcagc	5160
tccagccagg atggccacca gtggaccctc ttcttccaga atggcaaggc caagggtgtc	5220
cagggcaacc aggacagctt caccctgtg gtgaacagcc tggaccccc cctcctgacc	5280
agatacctga ggattcacc ccagagctgg gtccaccaga ttgcctgag gatggaggtc	5340
ctgggatgtg aggccaggga cctgtactga tgagcggccg ctcttagtag cagtatcgat	5400
aataaaagat ctttattttc attagatctg tgtgttggtt tttgtgtgt taattaagct	5460
cgcaaggaa ccctagtga tggagttggc cactccctct ctgcgcgctc gctcgctcac	5520
tgaggccggg cgaccaaagg tcgcccgcg ccggtgttt gcccgggcg cctcagttag	5580
cgagcgagcg cgagagagg gagtggccaa gacgatttaa atgacaagct tggcgtaatc	5640
atggtcatag ctgtttctg tgtgaaattg ttatccgctc acaattccac acaacatacg	5700
agccggaagc ataaagtgt aagcctgggg tgctaataga gtgagctaac tcacattaat	5760
tgcgttgccg tcaactgccc ctttccagtc gggaaacctg tcgtgccagc tgcattaatg	5820
aatcgcccaa cgcgcgggga gaggcggtt gcgtattggg cgctcttccg ctctcctgct	5880
cactgactcg ctgcgctcg tcgttcgggt gggcgagcg gtatcagctc actcaaaggc	5940
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg	6000
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg	6060
ccccctgac gagcatcaca aaaatcgac ctcaagtcag aggtggcgaa acccgacagg	6120
actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac	6180
cctgccgctt accggatacc tgcgcgcctt tctcccttcg ggaagcgtgg cgctttctca	6240
tagctcaagc tgtaggtatc tcagttcggg ttaggtcggt cgctccaagc tgggctgtgt	6300
gcacgaaccc ccgcttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc	6360

caaccocggtg agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag	6420
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac	6480
tagaagaaca gtatttggtg tctgcgctct gctgaagcca gttacottcg gaaaaagagt	6540
tggtagctct tgatccggca aacaaaccac cgttggtagc ggtgggtttt ttgtttgcaa	6600
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cttttgatct tttctacggg	6660
gtctgacgct cagtggaaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa	6720
aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat	6780
atatgagtaa acttgggtctg acagttacca atgcttaatc agtgaggcac ctatctcagc	6840
gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat	6900
acggggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc	6960
ggctccagat ttatcagcaa taaaccagcc agccygaagg gccgagcgca gaagtggctc	7020
tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta gagtaagtag	7080
ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg ttggtgtcacg	7140
ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg	7200
atcccccatg ttgtgcaaaa aagcggttag ctcttcggt cctccgatcg ttgtcagaag	7260
taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt	7320
catgccatcc gtaagatgct tttctgtgac ttggtgagta tcaaccaagt cattctgaga	7380
atagtgtatg cggcgaccga gttgctcttg cccggcgta atacgggata ataccgcgcc	7440
acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc	7500
aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac ccaactgac	7560
ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc	7620
cgcaaaaaag ggaataaggc cgacacggaa atgttgaata ctcatactct tcttttttca	7680
atattattga agcatttatc agggttattg tctcatgagc ggatacatat ttgaatgtat	7740
ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc cacctgacgt	7800
ctaagaaacc attattatca tgacattaac ctataaaaaat aggcgtatca cgaggccctt	7860
tcgtc	7865