(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) **Cộng hòa xã hội chủ nghĩa Việt Nam (VN)**
CỤC SỞ HỮU TRÍ TUỆ

(11) **1-0053115**

(51)[2020.01] **A61K 38/17**; A61P 29/00

(13) **B**

(21) 1-2022-02114
(22) 04/09/2020
(86) PCT/IB2020/058250 04/09/2020
(87) WO2021/044360 11/03/2021
(30) 19196045.9 06/09/2019 EP
(45) 25/11/2025 452
(43) 27/06/2022 411A
(73) NOVARTIS AG (CH)
Lichtstrasse 35, 4056 Basel, Switzerland
(72) IRIGARAY, Sebastien (FR); KLEIN, Laurent (FR); SKEGRO, Darko (HR); VILLANI, Marco (CH); WELZENBACH, Karl (DE).
(74) Công ty TNHH Ban Ca (BANCA)

(54) PROTEIN DUNG HỢP TRỊ LIỆU ĐỂ TĂNG CƯỜNG QUÁ TRÌNH TIÊU HỦY TẾ BÀO CHẾT VÀ TẾ BÀO CHỦ TÁI TỔ HỢP THÍCH HỢP ĐỂ SẢN XUẤT PROTEIN NÀY

(21) 1-2022-02114

(57) Sáng chế đề cập đến protein dung hợp trị liệu để tăng cường quá trình tiêu hủy tế bào chết bao gồm miền liên kết integrin, miền liên kết phosphatidylserin (PS) và miền hòa tan, trong đó miền hòa tan được chèn ở giữa miền liên kết integrin và miền liên kết PS; và trong đó miền liên kết integrin liên kết với integrin, và trong đó miền hòa tan là HSA và có trình tự axit amin của SEQ ID NO: 4, hoặc tương đồng trình tự ít nhất 90% với nó. Ngoài ra, sáng chế đề cập đến axit nucleic được phân lập mã hóa trình tự axit amin đã nêu; vectơ tách dòng hoặc biểu hiện chứa axit nucleic này; và tế bào chủ tái tổ hợp thích hợp để sản xuất protein dung hợp trị liệu đã nêu.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các protein dung hợp bao gồm cả khả năng liên kết integrin và liên kết phosphatidylserin. Các protein dung hợp này có thể được sử dụng làm trị liệu, cụ thể để ngăn ngừa hoặc điều trị các rối loạn viêm cấp tính hoặc mạn tính và cơ quan có nguồn gốc từ hệ miễn dịch hoặc đông máu và các rối loạn vi mạch.

Tình trạng kỹ thuật của sáng chế

Các tổn thương cơ quan viêm cấp tính (AOI) từ lâu đã là các bệnh với tỉ lệ mắc bệnh, tử vong cao và nhu cầu y tế chưa được đáp ứng một cách đáng kể. Các AOI điển hình bao gồm nhồi máu cơ tim (MI) và đột quỵ mà xuất hiện ở 32,4 triệu bệnh nhân trên toàn cầu mỗi năm. Bệnh nhân có MI và đột quỵ trước đó được Tổ chức Y tế Thế giới coi là nhóm nguy cơ cao đối với các biến cố liên quan đến mạch vành và não khác, mà xếp hạng trong số các nguyên nhân hàng đầu của tỉ lệ mắc bệnh trong thế giới phát triển. AOI khác là tổn thương thận cấp (AKI), mà xảy ra ở khoảng 13,3 triệu người mỗi năm. Ở các nước thu nhập cao, tỉ lệ mắc AKI là 3-5/1000 và có liên quan đến tỉ lệ tử vong cao (14-46%) (Metha *et al.*, (2015) Lancet, 385(9987): 2616-43). Tương tự với MI và đột quỵ, những người sống sót sau AKI thường không hồi phục hoàn toàn và có nguy cơ tăng phát triển bệnh thận mạn tính hoặc bệnh thận giai đoạn cuối. Cho đến nay chưa có một thuốc nào được FDA phê chuẩn để ngăn ngừa hoặc điều trị AKI. Việc phát triển các phương pháp điều trị mới cho AKI đã chứng tỏ đang là thử thách, mà chưa có kết quả thành công nào từ các thử nghiệm lâm sàng cho đến nay. Điều này có khả năng là do sinh lý bệnh đa nhân tố và nhiều khía cạnh của AKI bao gồm chứng viêm, rối loạn chức năng vi mạch và các cơ chế bệnh độc tố gây hư thận gây ra bởi các tổn thương nhiễm khuẩn, thiếu máu cục bộ/tái đẩy máu và/hoặc độc tố gây hư thận. Các yếu tố điều khiển này có thể tác động đồng thời hoặc liên tiếp để gây ra hầu hết tổn thương tế bào

không chỉ ống thận mà cả cầu thận, làm mất dự trữ chức năng thận và cuối cùng là suy thận.

Một mẫu số chung của AOI là gia tăng sự chết tế bào do tổn thương mô, gia tăng tạo ra các mảnh tế bào và vi hạt tiền huyết khối/tiền viêm mà có thể xâm nhập vào hệ tuần hoàn và mô bị thương tổn. Sau khi thâm nhập vào mô của bạch cầu trung tính để bảo vệ chống lại nhiễm trùng, bạch cầu trung tính trải qua quá trình chết theo chương trình hoặc các dạng chết tế bào khác trong mô bị ảnh hưởng. Bạch cầu trung tính chứa các chất có hại, bao gồm các enzym phân giải protein và các mô hình phân tử liên quan đến nguy hiểm (DAMP) có thể thúc đẩy tổn thương mô vật chủ và gây viêm. Sự hấp thụ hiệu quả của các tế bào sắp chết sẽ kích hoạt các sự kiện truyền tín hiệu mà dẫn đến việc lập trình lại các đại thực bào (MΦ) theo hướng không viêm, kiểu hình tiền phân giải và giải phóng các chất trung gian quan trọng để phân giải và sửa chữa thành công các mô bị ảnh hưởng. Quá trình tái lập trình này gần đây được cho là do sự truyền tín hiệu trao đổi chất mà kích hoạt các phản ứng chống viêm thể thực bào ở đại thực bào (Zhang *et al.*, (2019) Cell Metabolism, 29(2): 443-56). Việc loại bỏ các mảnh vụn hoặc các tế bào già hoặc sắp chết theo cách không gây viêm được gọi là 'quá trình tiêu hủy tế bào chết'.

Tuy nhiên, trong trường hợp quá trình tiêu hủy tế bào chết bị chậm lại, các tế bào hoại tử có thể tích tụ và gây ra, ví dụ, các phản ứng viêm kích hoạt các xytokin tiền viêm (TNF-α) hoặc IL-10 ức chế miễn dịch bởi các đại thực bào (Greenlee-Wacker (2016) Immunol. Reviews, 273: 357-370). Hơn nữa, nếu các mảnh vụn và hạt tế bào không được loại bỏ một cách hiệu quả, chúng có thể gây ra các cục và khối tụ tế bào, như các cụm mảnh bạch cầu trung tính-tiểu cầu, vi huyết khối và/hoặc giải phóng các mẫu phân tử liên quan đến nguy hiểm (DAMP) như ATP, ADN, histon hoặc HMGB1. Hậu quả có thể bao gồm tắc vi mạch, rối loạn chức năng và viêm nhiễm vô trùng rõ rệt dẫn đến sự tiến triển của tổn thương mô, suy cơ quan nguyên phát và thứ phát hoặc sửa chữa không thành công.

Trong giai đoạn cấp tính của AOI, các con đường tiêu hủy tế bào chết xuất hiện được điều hòa một cách đáng kể. Chứng viêm hoặc phản ứng cấp tính với chấn thương (tín hiệu cơ học, thiếu oxy huyết, ứng suất oxy hóa, bức xạ, viêm và nhiễm trùng) ngăn chặn quá trình tiêu hủy tế bào chết hoặc thực bào hiệu quả bằng cách điều hòa giảm các protein liên kết phosphatidylserin (PS) chuyên dụng mà bao gồm các protein bắc cầu và các thụ thể thanh thải/tiêu hủy tế bào chết ở bề mặt tế bào. Một ví dụ cho sự mất chức

năng của một thụ thể tiêu hủy tế bào chết là sự phân hủy protein của các thụ thể họ TAM như Mer tyrosin kinaza (MerTK). MerTK là một protein màng tích hợp được biểu hiện một cách ưu tiên trên các tế bào thực bào, nơi nó hoạt động như protein truyền tín hiệu nhưng cũng thúc đẩy quá trình tiêu hủy tế bào chết (thông qua các protein như Gas6 hoặc Protein S) và ức chế sự truyền tín hiệu viêm. Sự phân cắt protein phân giải và giải phóng miền ngoại bào hòa tan của MerTK được tạo ra bởi metalloproteinaza ADAM17. Quá trình rụng này có thể làm giảm quá trình tiêu hủy tế bào chết của tế bào thực bào bằng cách tước bỏ MerTK bề mặt. Ngoài ra, miền ngoại bào được giải phóng cũng có thể ức chế quá trình tiêu hủy tế bào chết *in vitro* (Zhang *et al.,* (2015) J Mol Cell Cardiol., 87:171-9; Miller *et al.,* (2017) Clin Cancer Res., 23(3):623-629). Lượng Mer hòa tan trong huyết thanh/huyết tương tăng thường được quan sát thấy ở các bệnh viêm, bệnh ác tính hoặc tự miễn như bệnh thận do tiểu đường hoặc bệnh lupus ban đỏ hệ thống (SLE) và có thể đánh dấu mức độ nghiêm trọng của bệnh (Ochodnicky P (2017) Am J Pathol., 187(9):1971-1983; Wu *et al.,* (2011) Arthritis Res Ther. 13:R88). Ngoài ra, các protein cầu nối như protein yếu tố EGF- hạt cầu béo sữa 8 (MFG-E8) cũng được điều hòa giảm trong các bệnh viêm cấp tính và mãn tính nhất. Tương tự như Mer hòa tan, nồng độ MFG-E8 trong huyết thanh/huyết tương giảm có thể được phát hiện ở bệnh nhân bị MI hoặc bệnh nhân đau thắt ngực ổn định (Dai *et al*., (2016) World J Cardiol., 8(1): 1-23) và có thể đánh dấu mức độ nghiêm trọng của bệnh như được mô tả đối với bệnh phổi tắc nghẽn mãn tính (COPD; Zhang *et al*., (2015) *nêu trên*).

Sự phơi nhiễm phosphatidylserin (PS) trên các tế bào sắp chết là một tín hiệu chống viêm và ức chế miễn dịch được bảo tồn về mặt tiến hóa đối với các tế bào miễn dịch. Một số lượng lớn các mầm bệnh động vật có vú sử dụng sự hấp thu qua trung gian PS như một phần của quá trình nhiễm độc tế bào (Birge *et al.,* (2016) Cell Death Diff., 23(6): 962-78). Ví dụ, virut có thể liên kết trực tiếp với các thụ thể liên kết PS hoặc thông qua các protein như Gas6 (Morizono & Chen (2014) J Virol., 88 (8): 4275-90). Có thể sự bất hoạt của các con đường thanh thải nội sinh để đáp ứng với chấn thương thể hiện một phản ứng được phát triển theo tiến hóa để làm giảm hiệu quả của tác nhân lây nhiễm xâm nhập và chiếm đoạt tế bào sau khi bị thương và do đó loại bỏ phản ứng và sự phòng thủ miễn dịch của vật chủ. Do đó, sự điều biến giảm các con đường thanh thải sẽ cải thiện hiệu quả của các yếu tố tác động miễn dịch bẩm sinh và tập nhiễm để chống lại nhiễm trùng. Là hệ quả "hỏa hoạn hữu hảo", quá trình tiêu hủy tế bào chết có

thể bị ảnh hưởng tạm thời trong quá trình tổn thương cơ quan cấp tính và các biến chứng nêu trên trong AOI có thể xảy ra. Sự tích tụ của các tế bào sắp chết, mảnh vỡ và các MP tiền viêm và tiền huyết khối là dấu hiệu của AOI và đại diện cho các tác nhân kích hoạt chính gây viêm và tổn thương vi mạch. Đáng chú ý là sự tích tụ của các vi hạt tiền viêm và tiền huyết khối thường gặp trong các bệnh nặng có nhu cầu y tế cao và có thể góp phần vào tỷ lệ mắc bệnh của họ. Ví dụ về các chỉ thị như vậy là nhiễm trùng máu và ung thư (Yang *et al.,* (2016) Tumour Biol., 37(6): 7881-91; Zhao *et al.,* (2016) J Exp Clin Cancer Res., 35: 54; Muhsin-Sharafaldine *et al.,* (2017) Biochim Biophys Acta Gen Subj., 1861(2): 286-295; Ma *et al.,* (2017) Sci Rep., 7(1): 4978; Souza *et al.,* (2015) Kidney Int. 87(6): 1100-8). Các nỗ lực khám phá thuốc trước đây trong lĩnh vực này đã tập trung vào các protein liên kết PS, mà có thể dùng làm cơ sở cho thiết kế ứng viên thuốc như đã được đánh giá tổng quan bởi (Li *et al.,* (2013) Exp Opin Ther Targets, 17(11): 1275-1285).

Một tập hợp con của các protein liên kết PS cũng nhận biết và liên kết với các integrin, như αvβ3 và αvβ5, mà được biểu hiện trên nhiều loại tế bào bao gồm các thể thực bào. Các protein này có tác dụng làm cầu nối giữa các tế bào chết theo chương trình/sắp chết tiếp xúc với PS với các integrin, dẫn đến quá trình tiêu hủy tế bào chết (còn gọi là quá trình thực bào) bởi các đại thực bào và các thể thực bào không chuyên nghiệp. Một số protein bắc cầu cũng được điều hòa giảm trong các bệnh viêm cấp tính và mãn tính nhất. Việc sử dụng điều trị đối với các protein bắc cầu như vậy hoặc các phiên bản bị cắt cụt của chúng đã được gợi ý trước đây (WO2006122327 (nhiễm trùng máu), WO2009064448 (tổn thương cơ quan do thiếu máu cục bộ/tái đẩy máu), WO2012149254 (thiếu máu cục bộ ở não) The Feinstein Institute for Medical Research; WO2015025959 (nhồi máu cơ tim) Kyushu University & Tokyo Medical University; WO20150175512 (sự tiêu xương) University of Pennsylvania; WO2017018698 (xơ hóa mô) Korea University Research và Business Foundation và US20180334486 (xơ hóa mô) Nexel Co., Ltd.); tuy nhiên việc sử dụng các protein kiểu dại hoặc protein xuất hiện tự nhiên bị giới hạn bởi nhiều vấn đề. Ví dụ, MFG-E8 kiểu dại (wtMFG-E8) được coi là có khả năng phát triển kém, độ hòa tan thấp, và xuất hiện ở hiệu suất rất thấp khi được nuôi cấy ở các hệ biểu hiện tế bào. Công trình của Castellanos *et al.,* (2016) đã chứng tỏ rằng MFG-E8 được bổ sung ở tế bào côn trùng hoặc tế bào CHO ở dạng thể dung hợp Fc-IgG bị kết tụ hoàn toàn và có thể chỉ được tinh chế một cách hiệu quả bằng cách bổ

sung các chất tẩy rửa như Triton X-100 hoặc CHAPS (Castellanos *et al.*, (2016) Protein Exp. Pur., 124: 10-22).

Sự tiêu hủy tế bào sắp chết, mảnh vụn và các vi hạt bởi các protein bắc cầu, ví dụ MFG-E8, EDIL3, Gas6, có thể loại bỏ các nguyên nhân chính gây viêm và rối loạn chức năng vi mạch nặng và do đó ngăn chặn sự tiến triển của tổn thương mô và có khả năng phân giải chứng viêm. Do đó, phương pháp điều trị để thúc đẩy sự thanh thải các tế bào sắp chết trong quá trình AOI có thể được sử dụng để làm giảm hoặc ít nhất là làm giảm nhẹ bệnh lý của AOI và có thể có ý nghĩa ở các thiết lập bệnh khác nơi mà các tế bào sắp chết hoặc các vi hạt tiếp xúc với PS bị thanh thải một cách không hiệu quả. Như vậy, cần có chất trị liệu mà có thể được sử dụng để làm giảm sự tổn thương mô và chứng viêm và có các đặc tính sản xuất mong muốn để giải quyết nhu cầu y tế chưa được đáp ứng ở AOI.

Bản chất kỹ thuật của sáng chế

Trong sáng chế, chủ đơn đã tạo ra các protein dung hợp tái tổ hợp trị liệu dựa vào cấu trúc của các protein bắc cầu xuất hiện tự nhiên (ví dụ MFG-E8) mà không có các tính chất và các vấn đề sản xuất không mong muốn nêu trên của protein kiểu dại. Các protein dung hợp theo sáng chế chứa miền liên kết integrin, miền liên kết PS và miền hòa tan. Các protein dung hợp này vẫn giữ các chức năng sinh học chính của protein MFG-E8 kiểu dại, ví dụ, bằng việc tạo chức để bắc cầu các tế bào sắp chết tiếp xúc với PS, mảnh vụn tế bào và vi hạt với các thể thực bào và do đó kích hoạt quá trình tiêu hủy tế bào chết. Ngoài ra, các protein dung hợp trị liệu theo sáng chế có khả năng phát triển được cải thiện, cụ thể là độ dính giảm và độ hòa tan được cải thiện so với protein MFG-E8 kiểu dại (SEQ ID NO: 1). Hơn thế nữa, các protein dung hợp trị liệu này có độ phơi nhiễm huyết tương dài hơn và có năng suất cao hơn khi được biểu hiện ở hệ biểu hiện tế bào khi so sánh với protein MFG-E8 kiểu dại.

Được đề xuất ở đây là các protein dung hợp trị liệu để tăng cường quá trình tiêu hủy tế bào chết bao gồm miền liên kết integrin, miền liên kết phosphatidylserin (PS) và miền hòa tan.

Theo một số phương án, miền hòa tan của protein dung hợp được liên kết với miền liên kết integrin. Theo một số phương án, miền hòa tan được liên kết với miền liên kết PS. Theo một số phương án, miền hòa tan được liên kết với cả miền liên kết integrin

và miền liên kết PS, tức là được định vị ở giữa miền liên kết integrin và miền liên kết PS. Theo một số phương án, miền hòa tan được chèn vào trong miền liên kết integrin hoặc được chèn vào trong miền liên kết PS. Theo một phương án, protein dung hợp trị liệu này có cấu trúc từ đầu cùng N đến đầu cùng C: miền liên kết integrin-miền hòa tan-miền liên kết PS.

Theo một số phương án, miền liên kết integrin của protein dung hợp trị liệu này bao gồm môtip liên kết Arginin-Glyxin-axit Aspartic (RGD) và liên kết với (các) intergrin αvβ3 và/hoặc αvβ5 hoặc α8β1.

Theo một số phương án, miền hòa tan của protein dung hợp trị liệu được liên kết trực tiếp với miền liên kết integrin và/hoặc được liên kết với miền liên kết PS tức là được chèn vào giữa các miền đã nêu. Theo phương án thay thế, miền hòa tan được liên kết gián tiếp với miền liên kết integrin và/hoặc miền liên kết PS bằng cầu nối, như cầu nối bên ngoài. Theo một số phương án, miền hòa tan bao gồm albumin huyết thanh người (HSA), miền 3 của HSA (HSA D3) hoặc vùng Fc của IgG (Fc-IgG), hoặc biến thể chức năng của chúng.

Theo một số phương án, protein dung hợp trị liệu bao gồm đầu cùng C của miền liên kết integrin được liên kết với đầu cùng N của miền hòa tan, và đầu cùng C của miền hòa tan được liên kết với miền liên kết PS. Theo một số phương án, protein dung hợp trị liệu bao gồm cấu trúc chung EGF-HSA-C1-C2 trong đó EGF là miền liên kết integrin, giống EGF của MFG-E8, EDIL3 hoặc các protein khác bao gồm miền liên kết integrin như được liệt kê trong Bảng 1, và C1-C2 là miền liên kết PS được tìm thấy trong MFG-E8, EDIL3 hoặc các protein khác bao gồm miền liên kết PS như được liệt kê trong Bảng 2. Các ví dụ về các protein chứa cả miền liên kết integrin và miền liên kết PS, ví dụ, MFG-E8 (SEQ ID NO: 1) và EDIL3 (SEQ ID NO: 11), được liệt kê trong Bảng 3.

Theo một số phương án, miền liên kết integrin là miền giống EGF, ví dụ, có trình tự axit amin như được nêu trong SEQ ID NO: 2 hoặc axit amin có độ tương đồng trình tự ít nhất 90%, 95%, 96%, 97%, 98% hoặc 99% với nó, hoặc các biến thể bị cắt cụt của nó. Theo một phương án, miền giống EGF bao gồm miền giống EGF của MFG-E8 người hoặc biến thể chức năng của nó bao gồm một, hai, ba, bốn, năm, lên đến 10 sự cải biến axit amin. Theo một phương án, miền giống EGF bao gồm miền giống EGF của EDIL3

người hoặc biến thể chức năng của nó bao gồm một, hai, ba, bốn, năm, lên đến 10 sự cải biến axit amin.

Theo một số phương án, miền liên kết PS bao gồm hai phân miền discoidin C1-C2, ví dụ, miền liên kết PS của MFG-E8 người có trình tự axit amin như được nêu trong SEQ ID NO: 3 hoặc axit amin có độ tương đồng trình tự ít nhất 90%, 95%, 96%, 97%, 98% hoặc 99% với nó, hoặc các biến thể bị cắt cụt của nó. Theo một phương án, miền liên kết PS bao gồm miền liên kết PS của MFG-E8 người hoặc biến thể chức năng của nó bao gồm một, hai, ba, bốn, năm, lên đến 10 sự cải biến axit amin. Theo một phương án, miền liên kết PS bao gồm miền liên kết PS của EDIL3 người hoặc biến thể chức năng của nó bao gồm một, hai, ba, bốn, năm, lên đến 10 sự cải biến axit amin.

Theo một số phương án, miền hòa tan là HSA hoặc biến thể chức năng của nó, ví dụ, có trình tự axit amin như được nêu trong SEQ ID NO: 4 hoặc axit amin có độ tương đồng trình tự ít nhất 90%, 95%, 96%, 97%, 98% hoặc 99% với nó, hoặc các biến thể bị cắt cụt của nó. Theo một phương án, HSA bao gồm sự thế axit amin C34S mà hoạt động để làm giảm thiên hướng protein kết tụ, và có trình tự axit amin như được nêu trong SEQ ID NO: 5. Theo một số phương án, miền hòa tan bao gồm albumin huyết thanh người (HSA) hoặc biến thể chức năng của nó bao gồm một, hai, ba, bốn, năm, lên đến 10 sự cải biến axit amin, ví dụ, HSA C34S, hoặc biến thể bị cắt cụt của HSA, ví dụ, miền 3 của HSA (HSA D3) hoặc biến thể chức năng của nó. Theo phương án được ưu tiên, miền hòa tan là HSA C34S.

Theo phương án thay thế, miền hòa tan bao gồm vùng Fc của IgG (Fc-IgG), ví dụ vùng Fc của IgG1, IgG2, IgG3 hoặc IgG4 người hoặc biến thể chức năng của nó. Theo một phương án, miền hòa tan bao gồm vùng Fc của Fc-IgG1 người có trình tự axit amin như được nêu trong SEQ ID NO: 7 hoặc axit amin có độ tương đồng trình tự ít nhất 90%, 95%, 96%, 97%, 98% hoặc 99% với nó, hoặc các biến thể bị cắt cụt của nó. Theo một phương án, Fc-IgG1 bao gồm các sự thế axit amin D265A và P329A để làm giảm chức năng tác động Fc, và có trình tự axit amin như được nêu trong SEQ ID NO: 8. Theo phương án khác, Fc-IgG1 bao gồm sự thế axit amin T366W để tạo ra 'ụ' hoặc nó có thể bao gồm sự thế axit amin T366S, L368A, Y407V để tạo thành 'lỗ'. Ngoài ra, ụ Fc-IgG1 có thể bao gồm sự thế axit amin S354C và lỗ Fc-IgG1 có thể bao gồm sự thế axit amin Y349C sao cho sự ghép cặp trên cầu nối xystein được tạo thành. Ngoài sự cải biến ụ trong lỗ, Fc-IgG1 cũng có thể bao gồm sự thế D265A và P329A để làm giảm

chức năng tác động Fc. Theo một phương án, Fc-IgG1 có trình tự axit amin như được nêu trong SEQ ID NO: 9 hoặc 10.

Theo phương án được ưu tiên, protein dung hợp trị liệu bao gồm protein yếu tố EGF hạt cầu béo sữa 8 (MFG-E8) và miền hòa tan, trong đó MFG-E8 bao gồm miền liên kết integrin giống EGF (SEQ ID NO: 2) và các miền C1-C2 liên kết phosphatidylserin (SEQ ID NO: 3 hoặc SEQ ID NO: 76). MFG-E8 có thể bao gồm MFG-E8 người xuất hiện tự nhiên hoặc kiểu dại (SEQ ID NO: 1), hoặc MFGE-8 với SEQ ID NO: 75, hoặc biến thể chức năng của nó. Theo một phương án, miền hòa tan được liên kết với đầu cùng N hoặc đầu cùng C của MFG-E8. Theo một phương án, miền hòa tan được chèn vào giữa miền giống EGF và miền C1 hoặc giữa miền C1 và C2. Theo phương án được ưu tiên, miền hòa tan được liên kết với đầu cùng C của miền giống EGF và được liên kết với đầu cùng N của miền C1. Miền hòa tan có thể được liên kết trực tiếp hoặc gián tiếp với đầu cùng C của miền giống EGF và được liên kết trực tiếp hoặc gián tiếp với đầu cùng N của miền C1. Theo một số phương án, sự liên kết gián tiếp là bằng cầu nối bên ngoài, ví dụ cầu nối dựa vào glyxin-serin.

Theo một phương án, protein dung hợp trị liệu bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 42 (FP330). Theo một phương án, protein dung hợp trị liệu có thể bao gồm thẻ histidin (thẻ His; SEQ ID NO: 67) để giúp phát hiện và/hoặc tinh chế trong thử nghiệm phân tích đặc điểm và biểu hiện protein. Theo một phương án, protein dung hợp trị liệu có thẻ His đầu cùng C và bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 44 (FP278). Các protein dung hợp trị liệu FP278 và FP330 có chung cùng trình tự axit amin ngoại trừ việc bổ sung thẻ His vào FP278.

Theo một số phương án, protein dung hợp trị liệu bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 42 (FP330) hoặc axit amin có độ tương đồng trình tự ít nhất 90%, 95%, 96%, 97%, 98% hoặc 99% với nó, hoặc các biến thể bị cắt cụt của nó. Ví dụ, protein dung hợp trị liệu FP776 bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 48 và có độ tương đồng trình tự 97,7% với FP330 (SEQ ID NO: 42). Ví dụ, protein dung hợp trị liệu FP068 bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 46 và có độ tương đồng trình tự 98,3% với FP330 (SEQ ID NO: 42). Ví dụ, protein dung hợp trị liệu FP816 bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 58 và có độ tương đồng trình tự 98,5% với FP330 (SEQ ID NO: 42). Ví dụ, protein dung hợp trị liệu FP811 bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 54

và có độ tương đồng trình tự 99,0% với FP330 (SEQ ID NO: 42). Ví dụ, protein dung hợp trị liệu FP010 bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 56 và có độ tương đồng trình tự 99,5% với FP330 (SEQ ID NO: 42). Ví dụ, protein dung hợp trị liệu FP138 bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 52 và có độ tương đồng trình tự 99,8% với FP330 (SEQ ID NO: 42). Ví dụ, protein dung hợp trị liệu FP284 bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 50 và có độ tương đồng trình tự 99,9% với FP330 (SEQ ID NO: 42).

Theo một số phương án, và như được mô tả trong phần Ví dụ, các protein dung hợp trị liệu theo sáng chế hoạt động để thúc đẩy quá trình tiêu hủy tế bào chết bởi tế bào nội mô ở thử nghiệm tiêu hủy tế bào chết của tế bào Jurkat bởi tế bào nội mô người và khôi phục quá trình tiêu hủy tế bào chết nền bị ảnh hưởng và được tăng cường bởi các đại thực bào ở thử nghiệm quá trình tiêu hủy tế bào chết của bạch cầu trung tính bởi đại thực bào người; các protein dung hợp hoạt động làm giảm số lượng vi hạt huyết tương bởi sự thanh thải thử nghiệm quá trình tiêu hủy tế bào chết của vi hạt bởi nội mô người; và/hoặc các protein dung hợp cung cấp sự bảo vệ chống lại sự tổn thương đa cơ quan ở mô hình thiếu máu cục bộ thận cấp.

Cũng được bộc lộ ở đây là các phương pháp, sử dụng, chất phản ứng chẩn đoán, dược phẩm và kit sử dụng hoặc chứa các protein dung hợp trị liệu này. Cũng được đề xuất ở đây là các axit nucleic mã hóa các protein dung hợp được bộc lộ, các vectơ tách dòng và biểu hiện bao gồm các axit nucleic như vậy, tế bào chủ chứa axit nucleic như vậy, và quy trình sản xuất các protein dung hợp được bộc lộ bằng cách nuôi cấy các tế bào chủ này.

Mô tả vắn tắt các hình vẽ

Hình 1 thể hiện sự biểu diễn bằng sơ đồ của các ví dụ về các protein dung hợp trị liệu theo sáng chế. Miền hòa tan (được đánh dấu là 'SD') được liên kết ở đầu cùng C, đầu cùng N, hoặc giữa các miền EGF, C1 hoặc C2 của MFG-E8.

Hình 2 thể hiện nhiều gel protein SDS-PAGE của các protein dung hợp được biểu hiện trong tế bào HEK. Hình 2A: protein EGF-HSA-C1-C2 (FP330; SEQ ID NO: 42); Hình 2B: EGF-HSA-C1-C2 của protein EDIL3 (FP050; SEQ ID NO: 12); Hình 2C: protein EGF-Fc(KiH) C1-C2 không giảm và giảm (protein này là heterodime của FP071 (EGF-Fc(ụ)-C1-C2; SEQ ID NO: 18) với lỗ Fc-IgG1 (SEQ ID NO: 10)); Hình 2D:

protein EGF-HSA-C1 (FP260; SEQ ID NO: 34). Đối với mỗi hình trong số các Hình 2A, 2C và 2D, cột thứ nhất thể hiện chỉ thị chuẩn không được nhuộm protein Precision plus và cột thứ hai thể hiện protein dung hợp tương ứng. Đối với Hình 2B, cột thứ nhất thể hiện protein dung hợp và cột thứ hai thể hiện chỉ thị chuẩn không được nhuộm protein Precision plus. Hình 2E thể hiện các protein tái tổ hợp khác mà đã được sản xuất và tinh chế.

Hình 3 minh họa hiệu quả mất MFG-E8 kiểu dại (wt) so với protein dung hợp FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) trong quá trình xử lý thực hành. Hình 3A thể hiện sự mất hiệu quả đối với wtMFG-E8 ở thử nghiệm cạnh tranh L-α-phosphatidylserin khi các dung dịch pha loãng protein được tạo ra ở các đĩa polypropylen (ký hiệu: □) so với các dung dịch pha loãng được tạo ra ở các đĩa không liên kết (ký hiệu: ●). Ngược lại, Hình 3B thể hiện trên thực tế không mất hiệu quả đối với protein dung hợp FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) ở thử nghiệm cạnh tranh PS khi các dung dịch pha loãng protein được tạo ra ở các đĩa polypropylen (ký hiệu: □) so với các đĩa không liên kết (ký hiệu: ●).

Hình 4 thể hiện sự liên kết của các protein dung hợp với L-α-phosphatidylserin. Hình 4A thể hiện sự liên kết của FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) với L-α-phosphatidylserin được cố định và với phạm vi yếu hơn với phospholipit cardiolipin, theo cách phụ thuộc nồng độ. Hình 4B thể hiện liên kết của wtMFG-E8 người và nhiều protein dung hợp trị liệu: FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44), FP250 (EGF-HSA; SEQ ID NO: 32), FP260 (EGF-HSA-C1; SEQ ID NO: 34), và FP270 (EGF-HSA-C2; SEQ ID NO: 36), với L-α-phosphatidylserin được cố định theo cách phụ thuộc nồng độ ở định dạng thử nghiệm cạnh tranh (sự cạnh tranh chống lại liên kết của wtMFG-E8 chuột được biotinyl hóa với L-α-phosphatidylserin).

Hình 5 thể hiện sự kết dính tế bào phụ thuộc αv-integrin với các protein dung hợp. Hình 5A chỉ ra rằng sự kết dính tế bào với FP330 (EGF-HSA-C1-C2; SEQ ID NO: 42) bị phong bế hoàn toàn bởi chất ức chế αv integrin cilengitide hoặc 10 mM EDTA. Đột biến điểm đơn ở môtip liên kết integrin RGD (RGD > RGE) của miền giống EGF (FP280; SEQ ID NO: 38) dẫn đến sự loại bỏ hoàn toàn sự kết dính tế bào như được thể hiện trên Hình 5B. Hình 5C chỉ ra rằng protein EGF-HSA được cố định (FP250; SEQ ID NO: 32) không hoặc chỉ hỗ trợ ở mức độ vừa phải sự kết dính của tế bào BW5147.G.1.4 bất kể miền giống EGF. Như được thể hiện trên Hình 5D, protein dung

hợp theo sáng chế này (FP330; SEQ ID NO: 42) thúc đẩy sự kết dính tế bào phục thuộc αv-integrin tương tự như wtMFG-E8 khi được biểu hiện ở tế bào CHO hoặc ở tế bào HEK.

Hình 6 thể hiện tác dụng của protein dung hợp trị liệu FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) đối với sự thúc đẩy quá trình tiêu hủy tế bào chết của các bạch cầu trung tính sắp chết bởi đại thực bào người. Nồng độ của protein dung hợp được thể hiện trên trục x và sự tiêu hủy tế bào chết [%] được thể hiện trên trục y.

Hình 7 thể hiện rằng protein dung hợp trị liệu FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) có thể giải cứu quá trình tiêu hủy tế bào chết bị ảnh hưởng bởi nội độc tố (lipopolysaccarit) bởi đại bào người. Hình 7A thể hiện sự suy giảm của quá trình tiêu hủy tế bào chết bởi các đại thực bào của các bạch cầu trung tính sắp chết của người là bằng 100 pg/ml lipopolysaccarit (LPS) ở ba đối tượng cho là người. Bảng bên trái thể hiện sự đáp ứng đối tượng cho riêng rẽ, bảng bên phải thể hiện sự suy yếu trung bình của sự tiêu hủy tế bào chết (%) của ba đối tượng cho. Hình 7B thể hiện sự giải cứu quá trình tiêu hủy tế bào chết bị ảnh hưởng bởi nội độc tố (LPS) của bạch cầu trung tính sắp chết bởi đại thực bào người khi có mặt protein dung hợp trị liệu FP278. Các chỉ số tiêu hủy tế bào chết của 3 đối tượng cho đại thực bào người khác nhau được chuẩn hóa và được lập biểu đồ ở dạng sự tiêu hủy tế bào chết (%).

Hình 8 thể hiện sự giải cứu sự suy yếu gây ra bởi hạt *S. aureus* của quá trình tiêu hủy tế bào chết của bạch cầu trung tính sắp chết bởi các đại thực bào người với protein dung hợp trị liệu FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44). Hình 8A thể hiện tác động của nồng độ của 100 nM FP278 đối với sự thúc đẩy quá trình tiêu hủy tế bào chết so với mức cơ sở (đường chấm chấm; phần bên tay trái của hình) cũng như tác dụng của 100 nM FP278 trong việc giải cứu sự làm suy yếu quá trình tiêu hủy tế bào chết gây ra bởi việc sử dụng *S. aureus* (phần bên tay phải của hình). Hình 8B thể hiện tác dụng của các nồng độ tăng dần của protein dung hợp FP278 (EC_{50} 8nM) đối với sự giải cứu quá trình tiêu hủy tế bào chết bị làm suy yếu gây ra bởi việc sử dụng *S. aureus*, và đối với sự thúc đẩy quá trình tiêu hủy tế bào chết khi các mức cơ sở của quá trình tiêu hủy tế bào chết đã đạt được.

Hình 9 thể hiện tác dụng của protein dung hợp trị liệu FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) đối với sự thúc đẩy quá trình tiêu hủy tế bào chết của tế

bào Jurkat sắp chết bởi tế bào nội mô người (HUVEC). Hiệu quả của protein dung hợp ở thử nghiệm quá trình tiêu hủy tế bào chết của tế bào nội mô phụ thuộc vào sự có mặt của miền nối tiếp C1-C2 hoặc C1-C1 khi, như được minh họa trên Hình 9, protein dung hợp có cấu trúc EGF-HSA-C2 (FP270; SEQ ID NO: 36) không có hiệu quả trong thử nghiệm này.

Hình 10 thể hiện rằng sự định vị của miền HSA ở protein dung hợp trị liệu, cụ thể là ở vị trí đầu cùng N đầu cùng C (FP220 (HSA-EGF-C1-C2; SEQ ID NO: 30) hoặc FP110 (EGF-C1-C2-HSA; SEQ ID NO: 28), tương ứng), đem lại chức năng phong bế quá trình tiêu hủy tế bào chết cho protein dung hợp MFG-E8 HSA ở thử nghiệm tiêu hủy tế bào chết của đại thực bào. Nồng độ của protein dung hợp được thể hiện trên trục x, sự tiêu hủy tế bào chết [%] được thể hiện trên trục y.

Hình 11 thể hiện phép so sánh sự thúc đẩy quá trình tiêu hủy tế bào chết bởi các định dạng khác nhau của các protein dung hợp trị liệu bao gồm HSA hoặc gốc Fc. Nồng độ của protein dung hợp được thể hiện trên trục x (nM), sự tiêu hủy tế bào chết [MFI] được thể hiện trên trục y. Hình 11A thể hiện sự so sánh các protein dung hợp bao gồm HSA với HSA được định vị ở đầu cùng C hoặc đầu cùng N hoặc giữa các miền giống EGF và miền C1; FP110 (EGF-C1-C2-HSA; SEQ ID NO: 28), FP220 (HSA-EGF-C1-C2; SEQ ID NO: 30) và FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44), tương ứng. Hình 11B thể hiện sự so sánh các protein dung hợp bao gồm gốc Fc với Fc được định vị ở đầu cùng C (FP060 (EGF-C1-C2-Fc [S354C,T366W]; SEQ ID NO: 14,) và FP080 (EGF-C1-C2-Fc; SEQ ID NO: 22)) hoặc ở giữa miền giống EGF và miền C1 (FP070 (EGF-Fc-C1-C2; SEQ ID NO: 16)) so với MFG-EG kiểu dại (SEQ ID NO: 1). Hai định dạng của gốc Fc được thể hiện: Fc kiểu dại (FP080; SEQ ID NO: 22) và gốc Fc với các cải biến S354C và T366W (cách đánh số châu Âu; FP060; SEQ ID NO: 14). Hình 11C thể hiện sự so sánh ba mẻ protein dung hợp FP090 (Fc-EGF-C1-C2; SEQ ID NO: 24) bao gồm gốc Fc được định vị ở đầu cùng N, ở ba nồng độ khác nhau (0,72, 7,2 và 72nM), so với đối chứng wt-MFG-E8. Hình 11D thể hiện sự thúc đẩy quá trình tiêu hủy tế bào chết bởi cấu trúc protein dung hợp FP050 bao gồm HSA được chèn vào giữa miền giống EGF và miền C1-C2 của EDIL3 (EGF-HSA-C1-C2 dựa vào EDIL3; SEQ ID NO: 12). Hình 11E thể hiện các ví dụ khác của các protein dung hợp theo sáng chế, ví dụ các biến thể khảm (FP145; SEQ ID NO: 80, FP1145; SEQ ID NO: 103, FP146; SEQ ID NO: 82, FP1146) và các tổ hợp của các miền liên kết integrin của MFGE8 hoặc

EDIL3 và miền liên kết PS như miền IgSF V của TIM4 hoặc miền GLA của protein bắc cầu GAS6 (FP1147 và FP1148). Hình 11F thể hiện chức năng thúc đẩy quá trình tiêu hủy tế bào chết của các protein dung hợp tái tổ hợp được cấu tạo là các ở khảm dung hợp các miền từ EDIL3 và MFG-E8 với thể chèn HSA. Số liệu này chỉ ra rằng FP145 (SEQ ID NO: 80) và FP146 (SEQ ID NO: 82) cảm ứng quá trình tiêu hủy tế bào chết của bạch cầu trung tính sắp chết bởi các đại thực bào người theo cách phụ thuộc nồng độ. Hình 11G thể hiện chức năng thúc đẩy quá trình tiêu hủy tế bào chết của các protein dung hợp tái tổ hợp được cấu tạo là các ở khảm dung hợp các miền từ EDIL3 và MFG-E8 với thể chèn HSA. Số liệu này chỉ ra rằng FP145 (SEQ ID NO: 80) và FP146 (SEQ ID NO: 82) cảm ứng quá trình tiêu hủy tế bào chết của tế bào Jurkat sắp chết bởi tế bào nội mô người (HUVEC) theo cách phụ thuộc nồng độ.

Hình 12 thể hiện sự thúc đẩy quá trình tiêu hủy tế bào chết bởi các tế bào HUVEC của protein dung hợp trị liệu FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) được thử nghiệm ở 3 nồng độ khác nhau lên đến 30 nM. Sự thúc đẩy quá trình tiêu hủy tế bào chết là phụ thuộc năng suất với quá trình tiêu hủy tế bào chết tăng khi nồng độ của protein dung hợp FP278 tăng.

Hình 13 thể hiện rằng các protein dung hợp trị liệu FP330 (EGF-HSA-C1-C2; SEQ ID NO: 42; Hình 13A), FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44; Hình 13B) và FP776 (EGF-HSA-C1-C2; SEQ ID NO: 48; Hình 13C) có thể giải cứu quá trình tiêu hủy tế bào chết bị ảnh hưởng bởi nội độc tố (lipopolysaccarit) bởi đại bào người. Nồng độ của protein dung hợp được thể hiện trên trục x, sự tiêu hủy tế bào chết [%] được thể hiện trên trục y.

Hình 14 thể hiện tác dụng của các protein dung hợp FP330 (EGF-HSA-C1-C2; SEQ ID NO: 42; Hình 14A), FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44; Hình 14B) và FP776 (EGF-HSA-C1-C2; SEQ ID NO: 48; Hình 14C) đối với sự thúc đẩy quá trình tiêu hủy tế bào chết của tế bào Jurkat sắp chết bởi tế bào nội mô người (HUVEC). Nồng độ của protein dung hợp được thể hiện trên trục x, sự tiêu hủy tế bào chết [%] được thể hiện trên trục y.

Hình 15 chỉ ra rằng liều đơn của các protein dung hợp trị liệu FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44), FP330 (EGF-HSA-C1-C2; SEQ ID NO: 42) hoặc FP776 (EGF-HSA-C1-C2; SEQ ID NO: 48) bảo vệ chức năng thận ở mô hình tổn

thương thận cấp (AKI) do tổn thương do thiếu máu cục bộ-tái tưới máu. Hình 15A chỉ ra rằng sự tăng creatinin huyết thanh (sCr) (mg/dL; trục y) bị giảm bởi sử dụng trong màng bụng (i.p.) 0,16mg/kg hoặc 0,5mg/kg FP278 (SEQ ID NO: 44) (trục x). Như được thể hiện trên Hình 15B, sử dụng trong tĩnh mạch (i.v.) 0,5mg/kg hoặc 1,5mg/kg protein dung hợp FP330 (SEQ ID NO: 42) làm giảm mức creatinin huyết thanh một cách đáng kể. Hình 15C thể hiện rằng việc sử dụng i.v. protein dung hợp FP776 (SEQ ID NO: 48) làm giảm creatinin huyết thanh theo cách phụ thuộc liều.

Hình 16 chỉ ra rằng liều đơn của protein dung hợp trị liệu FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) là 0,16mg/kg hoặc 0,5mg/kg, làm giảm mức ure nitơ máu (BUN) ở mô hình chuột bị tổn thương thận cấp.

Hình 17 chỉ ra rằng liều đơn của protein dung hợp trị liệu FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) bảo vệ các cơ quan xa khỏi sự đáp ứng pha cấp tính bị gây ra bởi AKI được cảm ứng bởi tái đẩy máu do thiếu máu cục bộ, dựa vào sự biểu hiện gen của các chỉ thị cho sự thương tổn. Hình 17A minh họa ví dụ sự đáp ứng gây ra bởi AKI như vậy của protein dạng tinh bột huyết thanh (SAA) ở tim chuột và Hình 17B minh họa ví dụ sự đáp ứng được cảm ứng bởi AKI (SAA) như vậy ở phổi chuột, cả hai chúng đều bị phong bế một cách tiềm năng sau một lần tiêm i.p. đơn lẻ của protein dung hợp có nguồn gốc từ MFGE8, FP278 (SEQ ID NO: 44) ở 0,16mg/kg hoặc 0,5 mg/kg/i.p.

Hình 18 thể hiện sự hấp thụ chất cản quang oxit sắt siêu thuận từ (SPIO) (Endorem®) bởi gan theo thời gian. Endorem® được tiêm trong tĩnh mạch ở tiêm bolus trong 1,2 giây vào con vật bị AKI (ở 24 giờ sau khi cảm ứng bệnh) hoặc sau sự hoạt động Giả bộ (các con vật sau thủ thuật cắt bỏ thận 24 giờ). Các con vật bị AKI thể hiện sự hấp thụ giảm đáng kể chất cản quang bởi gan (đích = tế bào Kupffer) so với các con vật Giả bộ. Điều trị với protein dung hợp FP776 (EGF-HSA-C1-C2; SEQ ID NO: 48) liều dự phòng -30 phút trước khi cảm ứng AKI, hoặc liều điều trị +5 giờ sau khi cảm ứng tổn thương tái tưới máu do chứng thiếu máu cục bộ, bảo vệ khỏi sự tích tụ chất cản quang ở gan của chuột nhắt AKI.

Mô tả chi tiết sáng chế

Được bộc lộ ở đây là các protein dung hợp trị liệu bao gồm miền liên kết integrin, miền liên kết PS và miền hòa tan. Cũng được bộc lộ ở đây là phương pháp điều trị sử

dụng các protein dung hợp theo sáng chế cũng như các thử nghiệm, như thử nghiệm tiêu hủy tế bào chết, hữu dụng để phân tích đặc điểm các protein dung hợp.

Định nghĩa

Để có thể hiểu rõ ràng sáng chế, các thuật ngữ nhất định được định nghĩa cụ thể trong toàn bộ phần mô tả chi tiết. Trừ khi có chỉ dẫn khác, tất cả các thuật ngữ khoa học và kỹ thuật dùng trong bản mô tả này có cùng nghĩa như thường được hiểu bởi người có hiểu biết trung bình trong lĩnh vực mà bản mô tả này thuộc về.

Trong tất cả các trường hợp mà thuật ngữ 'bao gồm', 'bao gồm có', 'chứa' hoặc dạng tương tự được sử dụng tham chiếu đến trình tự (ví dụ, trình tự axit amin), sẽ được hiểu là trình tự này cũng có thể bị giới hạn bởi thuật ngữ 'gồm', 'có', 'gồm có' hoặc dạng tương tự. Như dùng ở đây, cụm từ 'về cơ bản gồm' dùng để chỉ loại hoặc dạng của dược chất có hoạt tính được bao gồm trong phương pháp hoặc chế phẩm, cũng như là tá dược bất kỳ không có hoạt tính đối với mục đích dự kiến của phương pháp hoặc chế phẩm. Theo một số khía cạnh, cụm từ 'về cơ bản gồm' loại trừ rõ ràng sự bao gồm của một hoặc nhiều chất hoạt tính khác không phải là phân tử liên kết đa đặc hiệu theo sáng chế. Theo một số khía cạnh, cụm từ 'về cơ bản gồm' loại trừ rõ ràng sự bao gồm của một hoặc nhiều chất hoạt tính khác không phải là phân tử liên kết đa đặc hiệu theo sáng chế và chất đồng sử dụng thứ hai.

Thuật ngữ 'quá trình tiêu hủy tế bào chết' như được sử dụng ở đây để chỉ quá trình trong sinh học tế bào, trong đó tế bào sắp chết hoặc đã chết, như các tế bào chết theo chương trình hoặc hoại tử hoặc già hóa hoặc tế bào bị hoạt hóa cao hoặc các túi tế bào ngoại bào (vi hạt) hoặc mảnh vụn tế bào - được gọi chung là "mồi" - được loại bỏ bằng quá trình thực bào, tức là bị nhận chìm bởi tế bào thực bào và bị tiêu hủy. Trong quá trình tiêu hủy tế bào chết, các tế bào thực bào gắn chủ động và nhận chìm mồi này, tạo ra các túi được nạp đầy dịch lớn nội bào chứa mồi được gọi là efferosome, tạo thành khoang chứa lysosome nơi sự phân hủy mồi được bắt đầu. Trong quá trình chết theo chương trình, sự tiêu hủy tế bào chết đảm bảo rằng tế bào sắp chết được loại bỏ trước khi tính nguyên vẹn màng của chúng bị tổn hại và đồ chứa của chúng có thể rò rỉ vào trong mô xung quanh ngăn chặn sự tiếp xúc của mô xung quanh với DAMP như các enzym độc, chất oxy hóa và các thành phần nội bào khác như ADN, histon, và proteaza. Các tế bào thực bào chuyên nghiệp bao gồm các tế bào có nguồn gốc dòng tủy như đại

thực bào và tế bào đuôi gai nhưng không phải, ví dụ, tế bào mô đệm, cũng có thể thực hiện quá trình tiêu hủy tế bào chết như tế bào biểu mô và nội mô và nguyên bào sợi. Quá trình tiêu hủy tế bào chết bị suy giảm có thể được liên kết với các bệnh tự miễn và tổn thương mô và đã được chứng minh trong các bệnh như xơ nang, giãn phế quản, COPD, hen suyễn, xơ phổi vô căn, viêm khớp dạng thấp, lupus ban đỏ hệ thống, viêm cầu thận và xơ vữa động mạch (Vandivier RW *et al* (2006) Chest, 129(6): 1673-82). Không có liệu pháp nào đặc biệt thúc đẩy quá trình tiêu hủy tế bào chết đã được đưa vào sử dụng điều trị cho đến ngày nay.

Thuật ngữ 'thử nghiệm tiêu hủy tế bào chết' như được sử dụng ở đây và như được mô tả trong phần Ví dụ liên quan đến hệ thống xét nghiệm được phát triển để phân tích đặc tính của các protein dung hợp, mà sử dụng đại thực bào người hoặc tế bào nội mô người (HUVEC) làm tế bào thực bào. Được ví dụ ở đây là thử nghiệm tiêu hủy tế bào chết của đại thực bào-bạch cầu trung tính, thử nghiệm tiêu hủy tế bào chết của tế bào Jurkat bởi tế bào nội mô hoặc thử nghiệm tiêu hủy vi hạt của tế bào nội mô. Những thử nghiệm này, như được mô tả chi tiết hơn trong phần Ví dụ, có thể được sử dụng để chứng minh rằng liệu pháp sinh học có nguồn gốc từ MFG-E8 như protein dung hợp theo sáng chế, thúc đẩy một cách hiệu quả quá trình tiêu hủy tế bào chết sắp chết và các vi hạt bởi đại thực bào hoặc tế bào nội mô. Hơn thế nữa, thử nghiệm bạch cầu trung tính bởi đại thực bào được mô tả thích hợp để chứng tỏ rằng các hợp chất như vậy theo sáng chế có thể thậm chí giải cứu quá trình tiêu hủy tế bào chết bị suy giảm bởi LPS hoặc *S.aureus* của tế bào sắp chết.

Các thuật ngữ 'polypeptit' và 'protein' được sử dụng thay thế lẫn nhau ở đây để chỉ polyme của các gốc axit amin. Các cụm từ này cũng áp dụng cho polyme axit amin trong đó một hoặc nhiều gốc axit amin là gốc giả hóa học nhân tạo của axit amin xuất hiện tự nhiên tương ứng, cũng như là polyme axit amin xuất hiện tự nhiên và polyme axit amin không xuất hiện tự nhiên. Trừ khi có chỉ dẫn khác, trình tự polypeptit cụ thể cũng ngầm bao hàm các biến thể đã cải biến bảo toàn của chúng.

Thuật ngữ 'độ dính' như được sử dụng ở đây liên quan đến protein theo sáng chế để chỉ kết quả của sự gấp nếp nhầm protein mà thúc đẩy sự tạo cụm hoặc khối tụ protein. Các tác dụng không mong muốn và không hoạt động này là kết quả của sự tương tác kỵ nước bề mặt.

Như được sử dụng ở đây, 'đầu cùng C' để chỉ axit amin đầu cùng carboxyl của chuỗi polypeptit có nhóm carboxyl tự do (-COOH). Như được sử dụng ở đây, 'đầu cùng N' để chỉ axit amin đầu cùng amino của chuỗi polypeptit có nhóm amin tự do (-NH2).

Như được sử dụng ở đây, thuật ngữ 'protein dung hợp' để chỉ protein bao gồm nhiều miền, mà không thể tạo thành protein tự nhiên hoàn toàn hoặc kiểu dại nhưng có thể bị giới hạn ở miền hoạt tính của protein toàn vẹn chịu trách nhiệm đối với sự liên kết với thụ thể tương ứng của bề mặt của tế bào. Các protein dung hợp có thể được tạo ra bằng cách sử dụng thiết kế protein tái tổ hợp, trong đó thuật ngữ 'protein tái tổ hợp' để chỉ protein mà đã được điều chế, được biểu hiện, được tạo ra, hoặc được phân lập bởi công nghệ ADN tái tổ hợp. Dung hợp nối tiếp, ví dụ, để chỉ kỹ thuật nhờ đó các protein hoặc miền protein mong muốn được kết nối đơn giam đầu này với đầu kia qua sự dung hợp của đầu cùng N hoặc đầu cùng C giữa các protein. Điều này cung cấp cấu trúc cầu linh hoạt cho phép đủ không gian giữa các đối tác dung hợp để đảm bảo sự gấp nếp thích hợp. Tuy nhiên, đầu cùng N hoặc C của peptit thường là những thành phần quan trọng trong việc thu được kiểu gấp nếp mong muốn cho protein tái tổ hợp, với tác dụng là làm đơn giản hóa sự kết hợp đầu này với đầu kia của các miền có thể không hiệu quả. Ngoài ra, quá trình chèn miền liên quan đến sự dung hợp của các miền protein liên tiếp bằng cách mã hóa các cấu trúc mong muốn thành chuỗi polypeptit đơn lẻ và đôi khi chèn một miền bên trong miền khác. Trong cả hai quy trình đã đề cập ở trên, các miền được 'liên kết trực tiếp' hoặc 'liên kết một cách trực tiếp'. Việc chèn miền thường khó thực hiện hơn so với dung hợp nối tiếp do khó tìm được vị trí nối thích hợp ở gen quan tâm.

Ngoài các kỹ thuật dung hợp liên kết trực tiếp đã nói ở trên, cầu nối bên ngoài có thể được sử dụng để duy trì chức năng của các miền protein trong protein dung hợp. Cầu nối như vậy, dùng để chỉ một đoạn axit amin mà nối miền protein với miền protein khác và ở đây được gọi là 'cầu nối gián tiếp'. Vì vậy, các miền được 'liên kết gián tiếp' hoặc 'được liên kết một cách gián tiếp'. Ví dụ, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này đánh giá cao rằng polypeptit mà cấu trúc của nó bao gồm hai hoặc nhiều miền chức năng hoặc tổ chức thường bao gồm một đoạn axit amin ở giữa các miền như vậy mà liên kết chúng với nhau. Cầu nối này cho phép các tương tác miền, củng cố độ ổn định và có thể giảm sự cản trở không gian, mà thường khiến chúng được ưu tiên để sử dụng trong thiết kế protein được thiết kế ngay cả khi đầu cùng N và C có thể được dung hợp. Theo một số phương án, cầu nối được đặc trưng ở chỗ nó có xu hướng không áp

dụng cấu trúc không gian ba chiều cứng nhắc mà cung cấp độ mềm dẻo cho polypeptit này. Nhiều loại cầu nối xuất hiện tự nhiên khác nhau đã được sử dụng trong các protein được thiết kế, ví dụ, vùng bản lề globulin miễn dịch, mà có chức năng là cầu nối ở nhiều protein điều trị tái tổ hợp, cụ thể là ở các cấu trúc kháng thể được thiết kế di truyền (Pack P *et al.*, (1995) J. Mol. Biol.,246: 28-34). Bên cạnh các cầu nối tự nhiên, tập hợp các cầu nối nhân tạo đã được nghĩ ra, mà có thể được phân chia thành ba nhóm: cầu nối linh hoạt, cứng chắc và có thể phân tách được *in vivo*. (Yu K *et al.*, (2015) Biotech. Advances, 33(1): 155-64; Chen X *et al.*, (2013) Ad. Drug Delivery Reviews, 65(10): 1357-69). Các trình tự cầu nối linh hoạt được sử dụng rộng rãi nhất là (Gly)n (Sabourin *et al.*, (2007) Yeast, 24: 39-45) và (Gly_4Ser)n (SEQ ID NO: 64) (Huston *et al.*, 1988, 85: 5879-83) trong đó chiều dài cầu nối có thể được điều chỉnh bởi số lượng lần lặp "n". Theo một số phương án, polypeptit chứa yếu tố cầu nối có cấu trúc tổng thể ở dạng chung là D1-cầu nối-D2, trong đó D1 và D2 có thể là giống hoặc khác nhau và thể hiện hai miền liên kết với nhau bởi cầu nối này. Theo một số phương án, cầu nối polypeptit có chiều dài ít nhất 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 axit amin hoặc nhiều hơn.

'Sự cải biến' hoặc 'đột biến' của gốc/vị trí axit amin, như được sử dụng ở đây, để chỉ sự thay đổi của trình tự axit amin bậc một khi so sánh với trình tự axit amin bắt đầu, trong đó sự thay đổi này là từ sự thay đổi trình tự bao gồm gốc/vị trí axit amin đã nêu. Ví dụ, những cải biến điển hình bao gồm sự thế của gốc này (hoặc ở vị trí này) với axit amin khác (ví dụ, sự thế bảo toàn hoặc không bảo toàn), sự chèn một hoặc nhiều axit amin liền kề với gốc/vị trí đã nêu, và sự loại bỏ gốc/vị trí đã nêu. "Sự thế" axit amin hoặc biến thể của nó, để chỉ sự thay thế gốc axit amin hiện có ở trình tự axit amin đã định trước (ban đầu) với gốc axit amin khác. Nói chung và tốt hơn là, sự cải biến dẫn đến sự thay đổi ở ít nhất một hoạt tính sinh hóa lý của polypeptit biến thể so với polypeptit bao gồm trình tự axit amin ban đầu (hoặc 'kiểu dại').

Thuật ngữ 'biến thể được cải biến bảo toàn' áp dụng cho cả trình tự axit amin và axit nucleic. Đối với trình tự axit nucleic cụ thể, các biến thể đã cải biến bảo toàn dùng để chỉ các axit nucleic mà mã hóa cho trình tự axit amin tương đồng hoặc về cơ bản tương đồng, hoặc khi axit nucleic không mã hóa cho trình tự axit amin, mã hóa cho trình tự về cơ bản tương đồng. Do sự thoái hóa của mã di truyền, số lượng lớn các axit nucleic

tương đồng về mặt chức năng mã hóa protein đã cho bất kỳ. Ví dụ, các bộ ba GCA, GCC, GCG và GCU đều mã hóa axit amin alanin. Do đó, ở mọi vị trí trong đó alanin được xác định bởi codon, codon này có thể được thay đổi thành codon tương ứng bất kỳ được mô tả mà không làm thay đổi polypeptit được mã hóa. Các biến đổi axit nucleic này là 'sự biến đổi câm,' mà là một loại của sự thay đổi cải biến bảo toàn. Mọi trình tự axit nucleic ở đây mà mã hóa polypeptit cũng mô tả mọi sự biến đổi câm có thể có của axit nucleic. Người có hiểu biết trung bình trong lĩnh vực sẽ nhận ra rằng mỗi codon trong axit nucleic (ngoại trừ AUG, mà thông thường là codon duy nhất đối với methionin, và TGG, mà thông thường là codon duy nhất đối với tryptophan) có thể được cải biến để thu được phân tử tương đồng về mặt chức năng. Theo đó, mỗi sự thay đổi câm của axit nucleic mà mã hóa polypeptit được ẩn chứa trong mỗi trình tự được mô tả.

Đối với trình tự polypeptit, 'các biến thể đã cải biến bảo toàn' bao gồm sự thế, sự làm khuyết hoặc sự bổ sung riêng lẻ đối với trình tự polypeptit mà dẫn đến sự thế của axit amin bằng axit amin tương tự về mặt hóa học. Bảng sự thế bảo toàn nêu các axit amin tương tự về mặt chức năng được biết trong lĩnh vực. Các biến thể đã cải biến bảo toàn này thêm vào và không loại trừ các biến thể đa hình, các thể tương đồng giữa các loài, và các alen. Tám nhóm sau đây chứa các axit amin mà là sự thế bảo toàn đối với một axit amin khác: 1) Alanin (A), Glyxin (G); 2) axit Aspartic (D), axit Glutamic (E); 3) Asparagin (N), Glutamin (Q); 4) Arginin (R), Lysin (K); 5) Isoleuxin (I), Leuxin (L), Methionin (M), Valin (V); 6) Phenylalanin (F), Tyrosin (Y), Tryptophan (W); 7) Serin (S), Threonin (T); và 8) Xystein (C), Methionin (M) (xem, ví dụ, Creighton, Proteins (1984)). Theo một số phương án, cụm từ 'sự cải biến trình tự bảo toàn' được sử dụng để chỉ sự cải biến axit amin mà không ảnh hưởng hoặc thay đổi đáng kể các đặc tính liên kết của miền liên kết của các protein được thiết kế theo sáng chế.

'Biến thể protein' hoặc 'biến thể của protein' như được đề cập ở đây, liên quan đến protein bao gồm sự biến đổi trong đó một hoặc nhiều, ví dụ, 2, 3, 4, 5, 6, 7, 8, 9, 10 axit amin đã được cải biến. 'Biến thể chức năng' của protein như được đề cập ở đây, liên quan đến biến thể protein bao gồm sự cải biến mà dẫn đến sự thay đổi trình tự axit amin nhưng không có sự thay đổi đối với tính chất tổng thể của protein hoặc chức năng của nó. 'Biến thể bị cắt cụt' của protein như được đề cập ở đây, liên quan đến phiên bản bị cắt ngắn của protein nhưng phiên bản bị cắt ngắn của protein này vẫn giữ chức năng của protein ban đầu. Để xác định xem biến thể chức năng hoặc biến thể bị cắt cụt có

không thay đổi về đặc tính hoặc chức năng tổng thể hay không, các protein biến thể này có thể được thử nghiệm dựa trên protein ban đầu có chiều dài đầy đủ hoặc không được cải biến về tác dụng của chúng trong các thử nghiệm số os như được mô tả trong sáng chế. Ví dụ, việc thúc đẩy quá trình tiêu hủy tế bào chết bởi tế bào nội mô ở thử nghiệm tiêu hủy tế bào chết của tế bào Jurkat bởi tế bào nội mô, khôi phục quá trình tiêu hủy tế bào chết bị suy giảm bởi đại thực bào ở thử nghiệm tiêu hủy tế bào chết của bạch cầu trung tính bởi đại thực bào của người, làm giảm số lượng vi hạt huyết tương bằng cách thanh thải ở thử nghiệm tiêu hủy tế bào chết của vi hạt bởi nội mô người, và/hoặc cung cấp sự bảo vệ chống lại sự tổn thương đa cơ quan trong mô hình thiếu máu cục bộ thận cấp tính.

Thuật ngữ 'phần trăm tương đồng' hoặc 'phần trăm tương đồng trình tự' trong ngữ cảnh của hai hoặc nhiều axit nucleic hoặc trình tự polypeptit, để chỉ hai hoặc nhiều trình tự hoặc trình tự con mà giống nhau. Hai trình tự là 'về cơ bản giống nhau' và thể hiện 'độ tương đồng trình tự' nếu hai trình tự có tỷ lệ phần trăm xác định các gốc axit amin hoặc nucleotit giống nhau (tức là, độ tương đồng ít nhất 60%, tùy ý tương đồng ít nhất 65%, 70%, 75% , 80%, 85%, 90%, 95%, 96%, 97%, 98% hoặc 99% trên một vùng cụ thể hoặc, khi không được chỉ định, trên toàn bộ trình tự), khi được so sánh và được căn thẳng hàng trình tự để có được sự tương ứng tối đa qua một cửa sổ so sánh, hoặc vùng được chỉ định, ví dụ, như được đo bằng cách sử dụng một trong số các thuật toán so sánh trình tự sau đây hoặc bằng cách căn thẳng hàng trình tự thủ công và kiểm tra bằng mắt. Tùy ý, độ tương đồng tồn tại trên vùng có chiều dài ít nhất khoảng 50 nucleotit (hoặc 10 axit amin), hoặc trên vùng có chiều dài từ 100 đến 500 hoặc 1000, hoặc 2000 hoặc 3000 nucleotit hoặc nhiều hơn, hoặc theo cách khác, 30 đến 200, hoặc 300, hoặc 500, hoặc 700 hoặc 800 hoặc 900 hoặc 1000 axit amin hoặc nhiều hơn.

Để so sánh trình tự, thông thường một trình tự đóng vai trò làm trình tự tham chiếu, mà trình tự kiểm tra được so sánh với nó. Khi sử dụng thuật toán so sánh trình tự, trình tự kiểm tra và trình tự tham chiếu được nhập vào máy vi tính, các tọa độ trình tự con được chỉ định, nếu cần, và các tham số chương trình thuật toán trình tự được chỉ định. Các tham số chương trình mặc định có thể được sử dụng, hoặc các tham số thay thế có thể được chỉ định. Sau đó thuật toán so sánh trình tự tính tỉ lệ phần trăm độ tương đồng trình tự đối với trình tự kiểm tra tương quan với trình tự tham chiếu, dựa trên các tham số chương trình.

Thuật ngữ 'cửa sổ so sánh', như được sử dụng ở đây, bao gồm việc đề cập đến đoạn gồm một vị trí bất kỳ trong số nhiều vị trí axit nucleic hoặc axit amin liền kề được chọn từ nhóm bao gồm từ 20 đến 600, thường là từ khoảng 50 đến khoảng 200, thông thường hơn là từ khoảng 100 đến khoảng 150 trong đó trình tự có thể được so sánh với trình tự tham chiếu của cùng số lượng của các vị trí liền kề sau khi hai trình tự được căn thẳng hàng tối ưu. Các phương pháp căn thẳng hàng trình tự để so sánh là đã biết trong lĩnh vực kỹ thuật. Việc căn trình tự tối ưu của các trình tự để so sánh có thể được tiến hành, ví dụ như, bởi thuật toán tương đồng cục bộ của Smith và Waterman (1970) Adv. Appl. Math. 2:482c, bằng thuật toán căn trình tự tương đồng của Needleman & Wunsch (1970) J. Mol. Biol., 48: 443, bằng cách tra cứu phương pháp độ tương tự của Pearson & Lipman (1988) PNAS USA, 85: 2444, bởi các việc thực hiện được máy tính hóa của các thuật toán này (GAP, BESTFIT, FASTA, và TFASTA trong gói phần mềm di truyền Wisconsin, Nhóm máy tính di truyền, 575 Science Dr., Madison, WI), hoặc bằng cách căn trình tự thủ công và kiểm tra trực quan (xem, ví dụ, Brent *et al.*, (2003) Current Protocols in Molecular Biology).

Hai ví dụ của các thuật toán mà thích hợp cho việc xác định phần trăm độ tương đồng trình tự và độ tương tự trình tự là thuật toán BLAST và BLAST 2.0, mà được mô tả trong tài liệu Altschul *et al.*, (1977) Nuc. Acids Res,. 25: 3389-3402; và Altschul *et al.*, (1990) J. Mol. Biol., 215: 403-410, tương ứng. Phần mềm để thực hiện phân tích BLAST có sẵn công khai thông qua Trung Tâm Thông Tin Công Nghệ Sinh Học Quốc Gia (National Center for Biotechnology Information).

Thuật toán BLAST cũng thực hiện phép phân tích thống kê về sự tương tự giữa hai trình tự (xem, ví dụ, Karlin & Altschul (1993) PNAS. USA, 90: 5873-5787). Một phép đo độ tương tự được đưa ra bởi thuật toán BLAST là tổng xác suất nhỏ nhất (P(N)), mà đưa ra chỉ dấu về xác suất mà do nó, độ phù hợp giữa hai trình tự nucleotit hoặc axit amin có thể ngẫu nhiên xảy ra. Ví dụ, axit nucleic được coi là tương tự với trình tự tham chiếu nếu khả năng xảy ra tổng xác suất nhỏ nhất khi so sánh axit nucleic kiểm tra với axit nucleic tham chiếu nhỏ hơn khoảng 0,2, tốt hơn là nhỏ hơn khoảng 0,01, và tốt nhất là nhỏ hơn khoảng 0,001.

Độ tương đồng theo tỉ lệ phần trăm giữa hai trình tự axit amin cũng có thể được xác định bằng cách sử dụng thuật toán của E. Meyers và W. Miller (Comput. Appl. Biosci. 4:11-17 (1988)) mà đã được kết hợp vào chương trình ALIGN (phiên bản 2.0),

bằng cách sử dụng bảng gốc khối lượng PAM120, điểm phạt chiều dài khoảng trống bằng 12 và điểm phạt khoảng trống bằng 4. Ngoài ra, độ tương đồng theo tỉ lệ phần trăm giữa hai trình tự axit amin có thể được xác định bằng cách sử dụng thuật toán Needleman & Wunsch (*nêu trên*) mà đã được kết hợp vào chương trình GAP trong gói phần mềm GCG (có sẵn ở www.gcg.com), bằng cách cách sử dụng ma trận Blossom 62 hoặc ma trận PAM250, và lượng khoảng trống bằng 16, 14, 12, 10, 8, 6, hoặc 4 và lượng độ dài bằng 1, 2, 3, 4, 5, hoặc 6.

Polypeptit thường là về cơ bản giống với polypeptit thứ hai, ví dụ, trong đó hai polypeptit chỉ khác nhau bởi sự thế bảo toàn. Chỉ dẫn khác của việc hai trình tự axit nucleic về cơ bản tương đồng là việc hai phân tử hoặc phần bổ sung của chúng lai hóa với nhau trong điều kiện nghiêm ngặt.

Thuật ngữ 'axit nucleic' được sử dụng thay thế lẫn nhau trong bản mô tả này với thuật ngữ 'polynucleotit,' và dùng để chỉ deoxyribonucleotit hoặc ribonucleotit và polyme của chúng ở dạng mạch đơn hoặc mạch kép. 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, xuất hiện tự nhiên, và không xuất hiện 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 chất tương tự này bao gồm, mà không bị giới hạn ở, phosphorothioat, phosphoramidat, metyl phosphonat, chiral-metyl phosphonat, 2-O-metyl ribonucleotit, peptit-axit nucleic (PNA).

Trừ khi có chỉ dẫn khác, trình tự axit nucleic cụ thể cũng hoàn toàn bao hàm các biến thể được cải biến bảo toàn của chúng (ví dụ, sự thế codon thoái hóa) và các trình tự bổ sung, cũng như là trình tự được chỉ ra rõ ràng. Cụ thể là, sự thế bộ ba mã hóa thoái hóa có thể đạt được bằng cách tạo ra trình tự trong đó vị trí thứ ba của một hoặc nhiều bộ ba mã hóa được chọn (hoặc tất cả) được thế bằng các gốc bazơ hỗn hợp và/hoặc deoxyinosin (Batzer *et al.*, (1991) Nucleic Acid Res., 19: 5081; Ohtsuka *et al.*, (1985) J Biol Chem., 260: 2605-2608; và Rossolini *et al.*, (1994) Mol Cell Probes, 8: 91-98). Như được sử dụng ở đây, thuật ngữ, 'trình tự nucleotit được tối ưu hóa' nghĩa là trình tự nucleotit đã được thay đổi để mã hóa trình tự axit amin sử dụng các codon mà được ưu tiên trong tế bào sản xuất, ví dụ tế bào Buồng trứng chuột đồng Trung Quốc (CHO). Trình tự nucleotit đã được tối ưu hóa được thiết kế để giữ lại toàn bộ trình tự axit amin ban đầu được mã hóa bởi trình tự nucleotit bắt đầu, mà còn được gọi là trình tự 'gốc'.

Theo các phương án cụ thể, các trình tự được tối ưu hóa ở đây được thiết kế để có các codon mà được ưu tiên ở các tế bào động vật có vú CHO.

Các protein dung hợp trị liệu

Các miền liên kết integrin

Integrin là các thụ thể xuyên màng mà thúc đẩy sự kết dính tế bào - chất nền ngoại bào (ECM). Dựa vào liên kết phối tử, các integrin kích hoạt các con đường tải nạp tín hiệu mà làm trung gian cho các tín hiệu tế bào như sự điều hòa chu kỳ tế bào, sự sắp xếp bộ xương tế bào nội bào, và sự di chuyển của các thụ thể mới đến màng tế bào (Giancotti & Ruoslahti (1999) Science, 285 (5430): 1028-32). Sự có mặt của các integrin cho phép sự đáp ứng nhanh và linh hoạt đối với các sự kiện ở bề mặt tế bào. Một số loại integrin hiện có, và một tế bào có thể có nhiều loại khác nhau trên bề mặt của nó. Các integrin có hai tiểu đơn vị: α (alpha) và β (beta), mà mỗi tiểu đơn vị thâm nhập vào màng sinh chất và có một số miền tương bào (Nermut MV *et al* (1988). EMBO J., 7 (13): 4093-9). Axit amin tính axit đặc trưng ở vị trí tương tác integrin của nhiều protein ECM, ví dụ là bộ phận của trình tự axit amin Arginin-Glyxin-axit Aspartic ('RGD' ở mã axit amin một chữ cái). Môtip RGD đã được phát hiện ở nhiều protein nền như fibronectin, fibrinogen, vitronectin và osteopontin và hỗ trợ quá trình kết dính tế bào. Môtip RGD này được phát hiện ở nhiều protein ở miền protein được bảo toàn đã biết là miền giống EGF, mà nguồn gốc tên của nó từ yếu tố sinh trưởng biểu bì khi nó được mô tả lần đầu tiên. Miền giống EGF là một trong những miền phổ biến nhất được phát hiện ở các protein ngoại bào (Hidai C (2018) Open Access J Trans Med Res., 2(2): 67-71) và một số ví dụ về các miền giống EGF mà chứa môtip RGD được liệt kê dưới đây trong Bảng 1.

Bảng 1: Các ví dụ về các protein bao gồm protein miền giống EGF chứa môtip integrin RGD

Chữ viết tắt	**UniProt KB**	**Tên**	**Sự tham chiếu**
EDIL3	O43854	Đoạn lặp giống EGF và miền discoidin 3	Schürpf T *et al.*, (2012)

MFG-E8	Q08431	Protein Yếu tố EGF hạt cầu béo sữa 8	Taylor MR *et al.*, (1997)
NRG1	Q02297	Neuregulin-1	Leguchi K *et al.*, (2010)
IGFBP-1	P08833	Protein liên kết yếu tố sinh trưởng giống insulin 1	Haywood NJ *et al.*, (2017)
P2Y2R	P41231	Thụ thể nucleotit P2Y2	Erb L *et al.*, (2001)

Thuật ngữ 'miền liên kết integrin' như được sử dụng ở đây để chỉ đoạn axit amin, hoặc miền protein, mà có chức năng liên kết với các integrin. Theo một phương án của sáng chế, 'miền liên kết integrin' như được sử dụng ở đây để chỉ đoạn axit amin, hoặc miền protein, mà có chức năng liên kết với các integrin và bao gồm môtip RGD. Theo một phương án của sáng chế, miền liên kết integrin là miền giống EGF từ MFG-E8 người có trình tự axit amin như được nêu trong SEQ ID NO: 2. Theo phương án thay thế của sáng chế, miền liên kết integrin là miền giống EGF từ EDIL3 người (trình tự bất kỳ trong số các trình tự sau đây: SEQ ID NO: 11, SEQ ID NO: 77, SEQ ID NO: 96, SEQ ID NO: 97, SEQ ID NO: 98, SEQ ID NO: 99, SEQ ID NO: 100, hoặc SEQ ID NO: 101); ví dụ, trong đó các miền giống EGF có thể được phát hiện bên trong đoạn axit amin 1-132 của SEQ ID NO: 11.

Thuật ngữ 'liên kết với các integrin' như được sử dụng ở đây để chỉ hoạt tính liên kết integrin. Hoạt tính liên kết integrin có thể được xác định bằng phương pháp đã biết rõ trong lĩnh vực kỹ thuật. Ví dụ, thử nghiệm kết dính integrin được mô tả trong phần Ví dụ, phần 3.2 trong đó sự kết dính của tế bào u lympho biểu hiện αvβ3 integrin được đánh dấu huỳnh quang với các protein dung hợp trị liệu theo sáng chế được xác định. Miền liên kết integrin được coi là có hoạt tính liên kết integrin nếu nó có hoạt tính liên kết integrin ít nhất là 10%, như ví dụ ít nhất 25%, ít nhất 50%, ít nhất 75%, tốt hơn nữa là ít nhất 80%, như ít nhất 90%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98% khi được quan sát đối với protein MFG-E8 người (SEQ ID NO:1) khi được thử nghiệm bởi cùng phương pháp xác định hoạt tính tương ứng, tốt hơn là khi được thử nghiệm bằng cách sử dụng thử nghiệm được mô tả trong phần Ví dụ, phần 3.2.

Các miền liên kết phosphatidylserin

'Phosphatidylserin' (PS), như được sử dụng ở đây, liên quan đến phospholipit, mà là bộ phận của màng tế bào. PS hầu như bị giới hạn ở lớp trong của màng tế bào, trong khi phosphatidylcholin và sphingomyelin được định vị ở lớp ngoài. Sự phân bố bất đối xứng của các phospholipit được duy trì bởi hoạt động của các flippaza (P4-ATPaza như ATP11A và 11C) ở màng sinh chất để chuyển vị chủ động PS từ lớp ngoài vào lớp trong. Sự tiếp xúc bề mặt tế bào của PS được quan sát thấy không chỉ ở tế bào chết theo chương trình mà còn ở các tế bào lympho được kích hoạt, tiểu cầu được kích hoạt, hồng cầu bị già hóa, và một số tế bào ung thư và vi hạt tương ứng (Sakuragi *et al.*, (2019) PNAS USA, 116(8): 2907-12). Sự tiếp xúc PS có thể là chỉ thị sinh học cho tình trạng bệnh tiền huyết khối, viêm hoặc thiếu máu cục bộ (Pasalic *et al.*, (2018) J Thromb Haemost., 16(6): 1198-2010; Ma *et al.*, (2017) *nêu trên*; Zhao *et al.*, (2016) *nêu trên*. PS có chức năng ở nhiều con đường truyền tín hiệu tế bào và là phospholipit thiết yếu ở sự đông máu nơi mà nó có thể hoạt động là sự hình thành chất tăng cường của các phức hợp tenaza (các yếu tố IXa, VIIIa và X) và prothrombinaza (các yếu tố Xa, Va và prothrombin) (Spronk *et al.*, (2014) Thromb Res. 133 (Suppl 1): S54-6). Có thể chức năng được hiểu nhiều nhất của PS hiện bên ngoài vẫn là chỉ thị 'ăn tôi' cho các tế bào thực bào như đại thực bào để nhận chìm các tế bào chết theo chương trình, các mảnh vỡ của tế bào hoặc các tế bào được hoạt hóa tiếp xúc PS. Thuật ngữ 'miền liên kết phosphatidylserin' hoặc 'miền liên kết PS' như được sử dụng ở đây để chỉ một đoạn axit amin, hoặc miền protein, có chức năng liên kết với PS. Ví dụ về các protein nội sinh với các miền liên kết PS có thể được tìm thấy trong Bảng 2 dưới đây.

Bảng 2: Ví dụ về các thụ thể/protein có các miền liên kết phosphatidylserin

Chữ viết tắt	**UniProt**	**Tên**	**Miền liên kết PS giả định**	**Sự tham chiếu**
EDIL3	O43854	Các đoạn lặp giống EGF và miền discoidin 3	Miền discoidin C1-C2	Dasgupta et al., (2012)

MFG-E8	Q08431	protein yếu tố EGF hạt cầu béo sữa 8, lactadherin	Miền discoidin C1-C2	Andersen et al., (2000)
BAI1	O14514	Chất ức chế sự hình thành mạch máu đặc hiệu não 1	các đoạn lặp thrombospondin loại 1	Park *et al.*, (2007)
TIM1	Q96D42	Protein chứa globulin miễn dịch tế bào T và miền muxin 1	miền IgSF-V	Kobayashi *et al.*, (2007)
TIM3	Q8TDQ0	Protein chứa globulin miễn dịch tế bào T và miền muxin 3	miền IgSF-V	Cao *et al.*, (2007)
TIM4	Q96H15	Protein chứa globulin miễn dịch tế bào T và miền muxin 4	miền IgSF-V	Kobayashi *et al.*, (2007)
Stab1/Stab2	Q9NY15/ Q8WW/Q8	Stabilin-1 và -2	doạn lặp miền giống EGF (EGFrps) ở vùng ngoại bào	Park SY *et al.*, (2009)
TLT2	Q5T2D2	Thụ thể kích hoạt được biểu hiện trên protein giống tế bào dạng tủy 2	miền IgSF	de Freitas *et al.*, (2012)
TREM2	Q9NZC2	Thụ thể kích hoạt được biểu hiện trên trên tế bào dạng tủy 2	miền IgSF-V	Takahashi *et al.*, (2005)

CD300a	Q9U6N4	phân tử CD300a	miền IgSF-V	Simhadri *et al., (2012)*
RAGE	Q15109	Thu thể đối với sản phẩm cuối của quá trình glycat hóa tiến triển		He *et al.*, (2011)
AxV	P08758	Annexin V		Ravanat *et al.*, (1992)
PSR		Thụ thể phosphatidylserin		Mo *et al.*, (2003)
CD36	P16671	Glycoprotein tiểu cầu 4,		Banesh *et al.*, (2018)
CD68	P34810	Thụ thể chấp tẩy tạp lớp D		Chistiakov *et al.*, (2017)

Theo một phương án của sáng chế, miền PS thu được từ MFG-E8 người có trình tự axit amin như được nêu trong SEQ ID NO: 3, hoặc SEQ ID NO: 76. Theo phương án thay thế của sáng chế, miền liên kết integrin là miền liên kết PS từ EDIL3 người (SEQ ID NO: 11), trong đó miền liên kết PS bao gồm các axit amin 135-453 của SEQ ID NO: 11.

Hoạt tính liên kết PS có thể được xác định bằng phương pháp đã biết rõ trong lĩnh vực kỹ thuật. Ví dụ, thử nghiệm liên kết PS được mô tả ở phần Ví dụ, phần 3.1, trong đó sự liên kết của các protein dung hợp theo sáng chế với PS được phủ trên đĩa vi phiếm được đánh giá bằng sự cạnh tranh chống lại liên kết của MFG-E8 chuột được biotinyl hóa. Theo sáng chế, miền liên kết PS được coi là có hoạt tính liên kết PS nếu nó có hoạt tính liên kết PS ít nhất 10%, như ví dụ ít nhất 25%, ít nhất 50%, ít nhất 75%, ít nhất 80%, tốt hơn là ít nhất 90%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98% khi được quan sát đối với protein MFG-E8 người được thể hiện trong SEQ ID NO:1 khi

được thử nghiệm bởi cùng phương pháp xác định hoạt tính tương ứng, tốt hơn là khi được thử nghiệm bằng cách sử dụng thử nghiệm được mô tả trong phần Ví dụ, phần 3.1.

Protein bắc cầu

Có nhiều protein nội sinh mà bao gồm cả miền liên kết integrin và miền liên kết PS. Ví dụ về 'protein bắc cầu' như vậy được trình bày trong Bảng 3 dưới đây.

Bảng 3: Protein bắc cầu chứa cả miền liên kết integrin và miền liên kết phosphatidylserin

Chữ viết tắt	**UniProt**	**Tên**	**Miền liên kết PS giả định**	**Thụ thể trên thể thực bào**	**Sự tham chiếu**
EDIL3 (DEL-1)	O43854	Các đoạn lặp giống EGF và miền discoidin 3	Miền discoidin C1-C2	các integrin (αv- β2)	Dasgupta *et al.*, (2012)
MFG-E8	Q08431	protein yếu tố EGF hạt cầu béo sữa 8, lactadherin	Miền discoidin C1-C2	các integrin (αvb3/b5 α8b1)	Andersen *et al.*, (2000)
Pros1	P07225	Protein S	miền axit γ-carboxyglutamic (Gla)	"yếu tố chống đông máu" Tyro3 và Mer	Stitt *et al.*, (1995)
Gas6	Q14393	Protein đặc hiệu kìm hãm sinh trưởng 6	Miền Gla	Tyro3, Mer và AXL	Stitt *et al.*, (1995)

Để có giá trị điều trị, sẽ hữu dụng nếu protein bắc cầu bao gồm miền liên kết integrin mà nhận biết các integrin trên các thể thực bào mà thường không nhạy cảm với sự phân cắt hoặc rụng protein phân giải như đã được quan sát thấy ở các thành viên họ

TAM hoặc các thụ thể liên kết PS khác. Một protein có miền liên kết PS và miền liên kết integrin, ví dụ, MFG-E8 hoặc trình tự tương tự của nó EDIL3/DEL1, đã được chứng minh là gây ra hiện tượng tiêu hủy tế bào chết *in vitro* và do đó có thể có giá trị điều trị như các chất cảm ứng quá trình tiêu hủy tế bào chết ở AOI. Ngược lại, ví dụ, protein GAS6 có thể không đặc biệt hiệu quả trong việc thúc đẩy quá trình tiêu hủy tế bào chết ở AOI vì thụ thể của nó trên các thể thực bào (MerTK) bị phân cắt bằng phân giải protein trong quá trình viêm và nhiễm trùng như đã nêu ở trên.

Một ví dụ về protein bắc cầu, như được liệt kê trong Bảng 3 ở trên, là MFG-E8, mà là một trong những protein chính được tìm thấy trong màng cầu béo sữa (MFGM). MFG-E8 được biểu hiện và được tiết bởi một số loại tế bào khác nhau (ví dụ tế bào biểu mô tuyến vú, tế bào mạch máu, tế bào biểu mô mào tinh, tế bào cơ trơn động mạch chủ, đại thực bào hoạt hóa, nội mạc tử cung được kích thích và tế bào đuôi gai chưa trưởng thành) và mô (ví dụ: Tim, phổi, tuyến vú, lá lách, ruột, gan, thận, não, máu và nội mô). Protein MFG-E8 còn được gọi với nhiều tên khác nhau như, lactadherin, BP47, các thành phần 15/16, MFGM, MGP57/53, glycoprotein PAS-6/PAS-7, protein thành tế bào SED1, protein bề mặt tinh trùng SP47, kháng nguyên biểu mô vú BA46, và O-axetyl GD3 gangliosit synthaza (AGS). Gen MFG-E8 được định vị trên nhiễm sắc thể 1 ở chuột cống, nhiễm sắc thể 7 ở chuột nhắt, và nhiễm sắc thể 15 ở người. Sự cắt nối thay thế của tiền mARN của MFG-E8 dẫn đến ba dạng đồng dạng của protein người và hai dạng mARN, các dạng biến thể dài và ngắn được biểu hiện trong các tuyến vú của chuột nhắt. Gen MFG-E8 của người (UniProtKB - Q08431) mã hóa một protein dài 387 gốc được xử lý để tạo thành nhiều sản phẩm protein. Trình tự axit amin của MFG-E8 người, mà bao gồm peptit tín hiệu (các gốc 1-23; được gạch chân), miền giống EGF (các gốc 24-67; được in nghiêng), miền C1 (các gốc 70-225; in đậm), và miền C2 (các gốc 230-387; in đậm và gạch chân), được đưa ra dưới đây:

MPRPRLLAAL CGALLCAPSL LVA*LDICSKN PCHNGGLCEE ISQEVRGDVF PSYTCTCLKG YAGNHCE*TKC **VEPLGLENGN IANSQIAASS VRVTFLGLQH WVPELARLNR AGMVNAWTPS SNDDNPWIQV NLLRRMWVTG VVTQGASRLA SHEYLKAFKV AYSLNGHEFD FIHDVNKKHK EFVGNWNKNA VHVNLFETPV EAQYVRLYPT SCHTACTLRF ELLGCELNGC ANPLGLKNNS IPDKQITASS SYKTWGLHLF SWNPSYARLD KQGNFNAWVA GSYGNDQWLQ VDLGSSKEVT**

GIITQGARNF GSVQFVASYK VAYSNDSANW TEYQDPRTGS SKIFPGNWDN

HSHKKNLFET PILARYVRIL PVAWHNRIAL RLELLGC (SEQ ID NO: 1).

MFG-E8 thiếu chứng năng xuyên màng mà MFGM có và do đó đóng vai trò là protein màng ngoại biên. MFG-E8 người gôm một miền giống EGF đầu cùng N (SEQ ID NO: 2) mà liên kết với các integrin αvβ3 và αvβ5 được biểu hiện trên thể thực bào và miền liên kết PS (SEQ ID NO: 3) bao gồm hai phân miền F5/8-discoidin (C1 và C2) mà liên kết với ái lực cao với các phospholipit tích điện âm. Liên kết integrin là kết quả của của môtip RGD được định vị ở các gốc 46-48 của MFG-E8 người (SEQ ID NO: 1). Các tế bào chết theo chương trình, mảnh vụn tế bào, các tế bào được siêu hoạt hóa và phần lớn các vi hạt (MP) tiếp xúc PS và là các đích của MFG-E8 mà, hoạt động dưới dạng phân tử bắc cầu, opsonin hóa các tế bào này và vi hạt và liên kết chúng với các integrin αvβ3 và αvβ5 trên các thể thực bào. Hoạt động bắc cầu này kích hoạt chương trình nhận chìm hiệu quả dẫn đến sự nội hóa của các tế bào, mảnh vụn và vi hạt. Các protein được phát hiện ở MFGM được bảo toàn cao trong toàn bộ các loài. Cấu trúc protein MFG-E8 thay đổi theo loài; tất cả các loài được biết đến gần đây chứa hai miền C nhưng khác nhau ở số lượng miền giống EGF. Ví dụ, protein MFG-E8 người chứa một miền giống EGF, trong khi MFG-E8 bò và MFG-E8 chuột (SEQ ID NO: 68) có hai miền giống EGF, và gà, ếch và cá ngựa vằn có ba miền giống EGF. Các miền của MFG-E8, đã được đề xuất trước đó là thành phần cấu thành của các trị liệu, cụ thể là các miền liên kết PS (Kooijmans *et al.*, (2018) Nanoscale, 10(5): 2413-2426) và các mảnh của MFG-E8 đã được mô tả để hoạt động ở các mô hình xơ hóa (Đơn sáng chế Mỹ số US2018/0334486).

Sự hấp thụ không viêm của tế bào sắp chết, mảnh vụn tế bào và vi hạt bởi các thể thực bào chuyên nghiệp và không chuyên đóng vai trò quan trọng ở nội cân bằng sau tổn thương mô (Greenlee-Wacker (2016) *nêu trên*). Tầm quan trọng của sự thanh thải thích hợp trở nên rõ rệt hơn thế nữa ở các mô hình di truyền mà chuột nhắt bị bất hoạt MFG-E8 thể hiện, ví dụ, số lượng các tế bào sắp chết (không được thanh thải) tăng ở mô, sự đáp ứng viêm tăng quá mức ở mô hình bệnh như nhiễm trùng máu sơ sinh, tự miễn dịch, sự hình thành mạch kém và việc làm liền vết thương bị suy giảm (Hanayama *et al.*, (2004) Science, 204(5474): 1147-50; Das *et al.*, (2016) J Immunol., 196(12): 5089-5100; Hansen *et al.*, (2017) J Pediatr Surg., 52(9): 1520-7).

Ngoài ra, MFG-E8 đã được chứng minh là sinh ra môi trường dung nạp miễn dịch bởi sự ức chế hoạt hóa và tăng sinh tế bào T, ức chế các quần thể con Th1, Th2, và Th17 trong khi làm tăng tập con tế bào T điều hòa (Treg). Thú vị là, Treg góp phần vào việc giải quyết chứng viêm bằng cách cảm ứng quá trình tiêu hủy tế bào chết bởi các đại thực bào (Proto *et al.,* (2018) Immunity, 49(4): 666-77). MFG-E8 đã được mô tả là thúc đẩy sự gắn kết của các mô thu được từ tế bào gốc qua hàng rào MHC (Tan *et al.,* (2015) Stem Cell Reports, 5(5): 741-752). MFG-E8 cũng có nhiều sử dụng dinh dưỡng, mà giúp thúc đẩy sự phát triển và bảo vệ mô chống lại các tác nhân lây nhiễm. Các glycoprotein như MFG-E8 là các chế phẩm hỗ trợ dinh dưỡng tăng cường sức khỏe tiềm năng cho ứng dụng thực phẩm và dược phẩm. MFG-E8 cũng có thể được kết hợp với chất dinh dưỡng khác (ví dụ các vi sinh vật có lợi, mixen protein nước sữa, axit alpha-hyroxyisocaproic, xitrullin, và các axit béo mạch nhánh).

Miền hòa tan

Như được mô tả ở đây, các protein dung hợp trị liệu theo sáng chế bao gồm miền liên kết integrin và miền liên kết PS. Ngoài ra, các protein dung hợp cũng bao gồm miền bổ sung mà đem lại nhiều tính chất mong muốn cho protein dung hợp. Miền bổ sung này, mà đã được gọi là 'miền hòa tan' cho các mục đích của đơn này, đem lại các tính chất sinh học được cải thiện như độ hòa tan tăng, độ kết tụ giảm và hoạt tính sinh học tăng. Kết quả là, protein dung hợp thể hiện các đặc tính dược động lực học mong muốn. Hơn thế nữa, sự có mặt của miền hòa tan cải thiện độ ổn định của protein dung hợp trị liệu và dẫn đến sự biểu hiện được cải thiện của protein dung hợp so với protein kiểu dại ở hệ biểu hiện tế bào như được thể hiện bởi sự tăng năng suất sau khi tinh chế.

Sự có mặt của miền hòa tan cũng có thể đem lại thời gian bán hủy được kéo dài cho protein dung hợp trị liệu. Ví dụ, nhiều dược chất protein được liên kết với polyetylen glycol (PEG), reCODE PEG, khung kháng thể, axit polysialic (PSA), hydroxyetyl tinh bột (HES), và các protein huyết thanh, như albumin, IgG và FcRn, để kéo dài các thời gian bán hủy huyết tương của chúng và đạt được các tác dụng trị liệu được tăng cường (Kim *et al.,* (2010) J Pharmacol Exp Ther., 334: 682-92; Weimer *et al.*; (2008) Thromb Haemost. 99: 659-67; Dumont *et al.,* (2006) BioDrugs, 20: 151-60; Schellenberger *et al.,* (2009) Nat Biotechnol., 27: 1186-90).

Theo một số phương án, miền hòa tan là protein albumin như albumin huyết thanh người (HSA; SEQ ID NO: 4) hoặc các biến thể của nó. Ví dụ, HSA bao gồm sự thế axit amin C34S để làm giảm thiên hướng kết tụ (SEQ ID NO: 5), hoặc các miền của HSA như HSA D3; (SEQ ID NO: 6). HSA có thời gian bán hủy huyết thanh rất dài do nhiều yếu tố bao gồm kích thước tương đối lớn mà làm giảm sự lọc qua thận và đặc tính liên kết thụ thể Fc mới sinh (FcRn) của nó nhờ đó tránh được sự phân hủy nội bào. Việc sử dụng các mảnh đầu cùng N của HSA cho sự dung hợp với các polypeptit cũng đã được đề xuất (ví dụ Đơn sáng chế EP399666). Do đó, việc dung hợp hoặc liên hợp về mặt di truyền hoặc hóa học các phân tử với albumin có thể làm ổn định hoặc kéo dài tuổi thọ, và/hoặc duy trì hoạt tính của phân tử trong các khoảng thời gian kéo dài trong dung dịch, *in vitro* và/hoặc *in vivo*. Các phương pháp khác liên quan đến sự dung hợp HSA có thể được tìm thấy, ví dụ, trong các đơn sáng chế quốc tế WO2001/077137 và WO2003/060071.

Theo một số phương án, miền hòa tan bao gồm miền Fc kháng thể như Fc-globulin miễn dịch G1 người (Fc-IgG1; SEQ ID NO: 7). Miền Fc cũng có thể được cải biến, ví dụ, bằng cách sử dụng các cải biến dựa vào ụ bên trong lỗ (KiH) để cải thiện sự heterodime hóa của Fc bằng cách đưa những sự thế axit amin bổ sung vào trong miền CH3 của Fc. Ví dụ, sự thế T366W để tạo ra 'ụ' trên một miền CH3 và những sự thế T366S, L368A và Y407V để tạo ra 'lỗ' trên miền CH3 khác (Merchant *et al* (1998) Nat. Biotechnol., 16(7): 677-81; IgG1 cách đánh số châu Âu). Các cải biến khác mà có thể được bao gồm trong miền Fc một mình hoặc được kết hợp với các cải biến để cải thiện sự heterodime hóa có thể bao gồm, ví dụ, những sự thế axit amin thành xystein để tạo ra liên kết xystein bổ sung, ví dụ S354C và/hoặc Y349C, và những sự thế axit amin để làm giảm hoặc loại bỏ sự liên kết với các thụ thể Fcγ và protein bổ sung C1q, để 'làm câm' chức năng tác động miễn dịch. Đột biến kép được gọi là 'LALA' (L234A cùng với L235A; cách đánh số châu Âu) dẫn đến các chức năng tác động bị giảm (Lund *et al*., (1992) Mol Immunol., 29: 53-9). Theo cách khác, đột biến kép 'DAPA' (D265A cùng với P329A; cách đánh số châu Âu) dẫn đến chức năng tác động bị giảm. Theo một phương án của theo sáng chế, miền Fc có thể bao gồm những sự thế axit amin D265A, P329A để làm câm Fc và/hoặc những sự thế axit amin KiH T366W (ụ) hoặc T366S, L368A và Y407V (lỗ). Theo một phương án, miền Fc thu được từ IgG1 người và bao gồm những sự thế axit amin D265A, P329A (SEQ ID NO: 8). Theo phương án khác,

miền Fc thu được từ IgG1 người và bao gồm những sự thế axit amin D265A, P329A, S354C và sự thế axit amin T366W (Fc-IgG1-ụ; SEQ ID NO: 9). Theo phương án khác, miền Fc thu được từ IgG1 người và bao gồm những sự thế axit amin D265A, P329A, Y349C và những sự thế axit amin T366S, L368A và Y407V (Fc-IgG1-lỗ; SEQ ID NO: 10).

Theo một số phương án, miền hòa tan bao gồm miền Fc kháng thể thu được từ IgA, IgD, IgE hoặc IgM của người.

Theo một số phương án, miền hòa tan bao gồm SUMO (Yếu tố cải biến giống Ubiquitin nhỏ), Ubiquitin, GST (Glutathion S-transferaza), hoặc các biến thể của chúng.

Sự liên kết và sự định hướng của các miền của các protein dung hợp trị liệu

Miền liên kết integrin, miền liên kết PS và miền hòa tan của các protein dung hợp theo sáng chế được liên kết. Như được sử dụng ở đây, thuật ngữ 'được liên kết' hoặc 'liên kết' để chỉ một miền của protein dung hợp đang được gắn, trực tiếp hoặc gián tiếp, với miền khác của protein dung hợp. Sự gắn trực tiếp là một dạng của mối liên kết, và được gọi ở đây là 'được dung hợp' hoặc 'sự dung hợp'. Việc sử dụng phân tử có dạng A-B-C là một ví dụ: miền A được liên kết trực tiếp với miền B và được liên kết trực tiếp với miền C. Như vậy, miền A cũng có thể được mô tả là đang được dung hợp với miền B mà được dung hợp với miền C. Một ví dụ khác là, miền A được liên kết trực tiếp với miền B và được liên kết gián tiếp với miền C. Như vậy, miền A cũng có thể được mô tả là đang được dung hợp với miền B mà liên kết gián tiếp bởi cầu nối bên trong với miền C.

Theo một số phương án, kết nối là kết nối trực tiếp và các miền do đó được dung hợp với nhau. Theo một số phương án, miền liên kết integrin được dung hợp với miền liên kết PS mà được dung hợp với miền hòa tan. Cụ thể là, miền liên kết PS (ví dụ các phân miền discoidin C1-C2) được dung hợp với đầu cùng C của miền liên kết integrin (ví dụ miền giống EGF) và được dung hợp với đầu cùng N của miền hòa tan (ví dụ HSA). Theo một số phương án, miền hòa tan được dung hợp với miền liên kết integrin mà được dung hợp với miền liên kết PS. Cụ thể là, miền liên kết integrin (ví dụ miền giống EGF) được dung hợp với đầu cùng C của miền hòa tan (ví dụ HSA) và được dung hợp với đầu cùng N của miền liên kết PS (ví dụ phân miền discoidin C1-C2). Theo một số phương án, miền liên kết integrin được dung hợp với miền liên kết PS bao gồm các

phân miền discoidin C1-C2 và miền hòa tan được chèn ở giữa phân miền discoidin C1-C2. Cụ thể là, đầu cùng C của miền liên kết integrin (ví dụ miền giống EGF) được dung hợp với đầu cùng N của phân miền discoidin C1 và đầu cùng C của phân miền discoidin C1 được dung hợp với đầu cùng N của miền hòa tan (ví dụ HSA) và đầu cùng C của miền hòa tan được dung hợp với đầu cùng N của phân miền discoidin C2. Theo phương án khác, miền liên kết integrin được dung hợp với miền hòa tan mà được dung hợp với miền liên kết PS. Cụ thể là, miền hòa tan (ví dụ HSA) được dung hợp với đầu cùng C của miền liên kết integrin (ví dụ miền giống EGF) và với đầu cùng N của miền liên kết PS (ví dụ các phân miền discoidin C1-C2). Theo một phương án, HSA được dung hợp với đầu cùng C của miền giống EGF và được dung hợp với đầu cùng N của miền discoidin C1.

Theo một số phương án, miền hòa tan (ví dụ HSA) được dung hợp ở giữa miền liên kết integrin và miền liên kết PS. Theo một số phương án, miền liên kết integrin được định vị ở đầu cùng N của protein dung hợp và miền liên kết PS được định vị ở đầu cùng C của protein dung hợp.

Theo một số phương án, protein dung hợp bao gồm vùng thứ nhất chứa miền liên kết integrin, ví dụ miền giống EGF, vùng thứ hai chứa miền hòa tan (ví dụ HSA), và vùng thứ ba chứa miền liên kết PS, ví dụ miền discoidin C1 và/hoặc C2. Theo một số phương án, miền liên kết integrin được định vị ở đầu cùng N của protein dung hợp và miền liên kết PS được định vị ở đầu cùng C của protein dung hợp.

Theo một số phương án, miền hòa tan (ví dụ HSA) được dung hợp ở giữa miền liên kết integrin và miền liên kết PS. Theo một số phương án, miền liên kết integrin được định vị ở đầu cùng N của protein dung hợp và miền liên kết PS được định vị ở đầu cùng C của protein dung hợp.

Theo một số phương án, protein dung hợp bao gồm vùng thứ nhất chứa miền liên kết integrin, ví dụ miền giống EGF, vùng thứ hai chứa miền hòa tan (ví dụ HSA hoặc Fc), và vùng thứ ba chứa miền liên kết PS, ví dụ miền discoidin C1 và/hoặc C2. Theo một số phương án, miền liên kết integrin được định vị ở đầu cùng N của protein dung hợp và miền liên kết PS được định vị ở đầu cùng C của protein dung hợp.

Theo một số phương án, miền hòa tan là HSA.

Theo một số phương án, miền hòa tan là kháng thể Fc-globulin miễn dịch G1 (Fc-IgG1; SEQ ID NO: 7).

Theo một số phương án, miền hòa tan (ví dụ HSA) là HSA bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 5, hoặc biến thể chức năng của nó.

Theo phương án được ưu tiên, HSA bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 5 được dung hợp với đầu cùng C của miền giống EGF của MFG-E8 và được dung hợp với đầu cùng N của miền liên kết PS của MFG-E8. Theo một phương án, protein dung hợp bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 46 (FP068). Theo một phương án, protein dung hợp bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 48 (FP776).

Theo phương án thay thế, HSA bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 5 được dung hợp với đầu cùng C của miền giống EGF của EDIL3 và được dung hợp với đầu cùng N của miền liên kết PS của EDIL3. Theo một phương án, protein dung hợp bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 70 (FP1068). Theo một phương án, protein dung hợp bao gồm trình tự axit amin như được nêu trong SEQ ID NO: 69 (FP1776).

Theo một số phương án, kết nối là qua cầu nối polypeptit và cầu nối polypeptit mà, ví dụ, nối miền hòa tan với miền liên kết PS ở protein dung hợp theo sáng chế được gọi là 'cầu nối bên ngoài'. Các cầu nối bên ngoài này thông thường bao gồm glyxin (G) và/hoặc serin (S) và cũng có thể bao gồm glyxin và leuxin (GL) hoặc glyxin và valin (GL). Theo một số phương án, cầu nối này bao gồm nhiều gốc G và S, ví dụ, G_2S và nhiều gốc này như $(G_2S)_4$ như được nêu trong SEQ ID NO: 62, $(GS)_4$ như được nêu trong SEQ ID NO: 63, G_4S như được nêu trong SEQ ID NO: 64 hoặc $(G_4S)_2$ như được nêu trong SEQ ID NO: 65.

Theo một số phương án, cầu nối bên ngoài được dung hợp giữa đầu cùng C của miền liên kết integrin và đầu cùng N của miền hòa tan. Cụ thể là, cầu nối bên ngoài được dung hợp với đầu cùng C của miền giống EGF và đầu cùng N của HSA. Theo một số phương án, cầu nối bên ngoài được dung hợp ở giữa đầu cùng C của miền hòa tan và đầu cùng N của miền liên kết PS. Cụ thể là cầu nối bên ngoài được dung hợp với đầu cùng C của HSA và đầu cùng N của miền liên kết PS. Theo một số phương án, cầu nối bên ngoài được dung hợp ở giữa đầu cùng C của miền liên kết integrin và đầu cùng N

của miền hòa tan, và cầu nối bên ngoài bổ sung được dung hợp ở giữa đầu cùng C của miền hòa tan và đầu cùng N của miền liên kết PS. Cụ thể là, cầu nối bên ngoài được dung hợp với đầu cùng C của miền giống EGF và đầu cùng N của HSA, và cầu nối bên ngoài bổ sung được dung hợp với đầu cùng C của HSA và đầu cùng N của miền liên kết PS.

Theo một số phương án, cầu nối bên ngoài bao gồm GS được dung hợp với đầu cùng C của miền liên kết integrin và với đầu cùng N của miền hòa tan. Theo một số phương án, cầu nối bên ngoài bao gồm GL được dung hợp với đầu cùng C của miền hòa tan và với đầu cùng N của miền liên kết PS. Theo một số phương án, cầu nối bên ngoài bao gồm $(G_2S)_4$ (SEQ ID NO: 62) được dung hợp với đầu cùng C của miền hòa tan và với đầu cùng N của miền liên kết PS. Theo một số phương án, cầu nối bên ngoài bao gồm G_4S (SEQ ID NO: 64) được dung hợp với đầu cùng C của miền hòa tan và với đầu cùng N của miền liên kết PS. Theo một số phương án, cầu nối bên ngoài bao gồm $(G_4S)_2$ (SEQ ID NO: 65) được dung hợp với đầu cùng C của miền hòa tan và với đầu cùng N của miền liên kết PS.

Theo một phương án, cầu nối bên ngoài bao gồm GS được dung hợp với đầu cùng C của miền giống EGF và với đầu cùng N của HSA. Protein dung hợp theo sáng chế bao gồm cấu trúc này có trình tự axit amin như được nêu trong SEQ ID NO: 42 (FP330).

Theo một phương án, cầu nối bên ngoài bao gồm GS được dung hợp với đầu cùng C của miền giống EGF và với đầu cùng N của HSA, và cầu nối bên ngoài khác bao gồm $(GS)_4$ (SEQ ID NO: 63) được dung hợp với đầu cùng C của HSA và với đầu cùng N của miền liên kết PS.

Theo một phương án, cầu nối bên ngoài bao gồm GS được dung hợp với đầu cùng C của miền giống EGF và với đầu cùng N của HSA, và cầu nối bên ngoài khác bao gồm $(G_2S)_4$ (SEQ ID NO: 62) được dung hợp với đầu cùng C của HSA và với đầu cùng N của miền liên kết PS. Protein dung hợp theo sáng chế bao gồm cấu trúc này có trình tự axit amin như được nêu trong SEQ ID NO: 42 (FP330).

Theo một phương án, cầu nối bên ngoài bao gồm GS được dung hợp với đầu cùng C của miền giống EGF và với đầu cùng N của HSA. Đầu cùng C của HSA được dung hợp trực tiếp với đầu cùng N của miền liên kết PS.

Theo một phương án, cầu nối bên ngoài bao gồm GS được dung hợp với đầu cùng C của miền giống EGF và với đầu cùng N của HSA, và cầu nối bên ngoài bổ sung bao gồm G_4S (SEQ ID NO: 64) được dung hợp với đầu cùng C của HSA và với đầu cùng N của miền liên kết PS. Protein dung hợp theo sáng chế bao gồm cấu trúc này có trình tự axit amin như được nêu trong SEQ ID NO: 54 (FP811).

Theo một phương án, cầu nối bên ngoài bao gồm GS được dung hợp với đầu cùng C của miền giống EGF và với đầu cùng N của HSA, và cầu nối bên ngoài khác bao gồm $(G_4S)_2$ (SEQ ID NO: 65) được dung hợp với đầu cùng C của HSA và với đầu cùng N của miền liên kết PS. Protein dung hợp theo sáng chế bao gồm cấu trúc này có trình tự axit amin như được nêu trong SEQ ID NO: 56 (FP010).

Theo một số phương án, thẻ His được dung hợp với cầu nối bên ngoài bao gồm GS (GS-6xHis; SEQ ID NO: 66) mà được dung hợp với đầu cùng C của miền liên kết PS. Theo một phương án, protein dung hợp theo sáng chế bao gồm thẻ His có trình tự axit amin như được nêu trong SEQ ID NO: 44 (FP278) hoặc SEQ ID NO: 60 (FP114 hoặc FP260).

Các tính chất chức năng của các protein dung hợp trị liệu

Sáng chế đề xuất các protein dung hợp có nguồn gốc từ MFG-E8 người và có hiệu quả trong việc thúc đẩy quá trình tiêu hủy tế bào chết và do đó có hoạt tính trong việc loại bỏ yếu tố điều khiển quan trọng của chứng viêm toàn thân và bệnh lý vi mạch. Như được nêu trong phần Ví dụ, các protein dung hợp có cấu trúc chung EGF-HSA-C1-C2 đã được chứng minh là có hiệu quả ở nhiều phân tích tiêu hủy tế bào chết. Ví dụ, các protein dung hợp đã có hiệu quả trong quá trình tiêu hủy tế bào chết bị làm suy yếu bởi việc lipopolysaccarit (LPS) dự trữ hoặc *S.aureus* của đại thực bào và tăng cường quá trình tiêu hủy tế bào chết của vi hạt và các tế bào sắp chết bởi các tế bào nội mô. Các protein dung hợp cũng đã có hiệu quả trong việc bảo vệ chức năng thận và bảo vệ chống lại việc giảm thể trọng ở mô hình chuột nhắt bị tổn thương thận cấp.

Các trình tự protein làm ví dụ

Các trình tự axit amin trong Bảng 4 bao gồm các ví dụ về các protein dung hợp trị liệu theo sáng chế, cũng như các phần của chúng.

Trong toàn bộ bản mô tả này, nếu có sự không thống nhất giữa nội dung của bản mô tả này (ví dụ, Bảng 4) và danh mục trình tự, thì nội dung của bản mô tả này sẽ có tính quyết định.

Bảng 4. Các trình tự protein làm ví dụ

SEQ ID NO	**Mô tả**	**Trình tự**
1	Người MFG-E8	MPRPRLLAALCGALLCAPSLLVALDICSKNPCHNGGLC EEISQEVRGDVFPSYTCTCLKGYAGNHCETKCVEPLGL ENGNIANSQIAASSVRVTFLGLQHWVPELARLNRAGM VNAWTPSSNDDNPWIQVNLLRRMWVTGVVTQGASRL ASHEYLKAFKVAYSLNGHEFDFIHDVNKKHKEFVGNW NKNAVHVNLFETPVEAQYVRLYPTSCHTACTLRFELLG CELNGCANPLGLKNNSIPDKQITASSSYKTWGLHLFSW NPSYARLDKQGNFNAWVAGSYGNDQWLQVDLGSSKE VTGIITQGARNFGSVQFVASYKVAYSNDSANWTEYQD PRTGSSKIFPGNWDNHSHKKNLFETPILARYVRILPVAW HNRIALRLELLGC
2	Miền giống EGF của MFG-E8	LDICSKNPCHNGGLCEEISQEV**RGD**VFPSYTCTCLKGY AGNHCETK
3	Miền liên kết PS của MFG-E8 (các phân miền C1-C2)	CVEPLGLENGNIANSQIAASSVRVTFLGLQHWVPELAR LNRAGMVNAWTPSSNDDNPWIQVNLLRRMWVTGVVT QGASRLASHEYLKAFKVAYSLNGHEFDFIHDVNKKHK EFVGNWNKNAVHVNLFETPVEAQYVRLYPTSCHTACT LRFELLGCELNGCANPLGLKNNSIPDKQITASSSYKTWG LHLFSWNPSYARLDKQGNFNAWVAGSYGNDQWLQVD LGSSKEVTGIITQGARNFGSVQFVASYKVAYSNDSANW

		TEYQDPRTGSSKIFPGNWDNHSHKKNLFETPILARYVRI LPVAWHNRIALRLELLGC
4	HSA kiểu dại	DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNL PRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFY APELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDE GKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPK AEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAK YICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPAD LPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTK KVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAE DYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCF SALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQT ALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADD KETCFAEEGKKLVAASQAAL
5	HSA (C34S)	DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQSPFED HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNL PRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFY APELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDE GKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPK AEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAK YICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPAD LPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTK KVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAE DYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCF

		SALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQT ALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADD KETCFAEEGKKLVAASQAAL
6	HSA D3	LVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQ VSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDCLS VFLNQLCVLHEKTPVSDRVTKCCTESLVNGRPCFSALE VDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVE LVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETC FAEEGKKLVAASQAALGL
7	Fc-IgG1 kiểu dại	APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHE DPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSV LTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQP REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEW ESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQ GNVFSCSVMHEALHNHYTQKSLSLSPGK
8	Fc-IgG1 bị câm	APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVAVSHE DPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSV LTVLHQDWLNGKEYKCKVSNKALAAPIEKTISKAKGQ PREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVE WESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRW QQGNVFSCSVMHEALHNHYTQKSLSLSPGK
9	Fc-IgG1 Ụ	APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVAVSHE DPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSV LTVLHQDWLNGKEYKCKVSNKALAAPIEKTISKAKGQ PREPQVYTLPPCREEMTKNQVSLWCLVKGFYPSDIAVE WESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRW QQGNVFSCSVMHEALHNHYTQKSLSLSPGK
10	Fc-IgG1 Lỗ	APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVAVSHE DPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSV

		LTVLHQDWLNGKEYKCKVSNKALAAPIEKTISKAKGQ PREPQVCTLPPSREEMTKNQVSLSCAVKGFYPSDIAVE WESNGQPENNYKTTPPVLDSDGSFFLVSKLTVDKSRW QQGNVFSCSVMHEALHNHYTQKSLSLSPGK
11	EDIL3 người	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA NYSCECPGEFMGRNCQYKCSGPLGIEGGIISNQQITASS THRALFGLQKWYPYYARLNKKGLINAWTAAENDRWP WIQINLQRKMRVTGVITQGAKRIGSPEYIKSYKIAYSND GKTWAMYKVKGTNEDMVFRGNIDNNTPYANSFTPPIK AQYVRLYPQVCRRHCTLRMELLGCELSGCSEPLGMKS GHIQDYQITASSIFRTLNMDMFTWEPRKARLDKQGKVN AWTSGHNDQSQWLQVDLLVPTKVTGIITQGAKDFGHV QFVGSYKLAYSNDGEHWTVYQDEKQRKDKVFQGNFD NDTHRKNVIDPPIYARHIRILPWSWYGRITLRSELLGCT EEE
12	FP050 EDIL3 EGF-HSA-C1-C2	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA NYSCECPGEFMGRNCQYKGSDAHKSEVAHRFKDLGEE NFKALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVA DESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCA KQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHD NEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTEC CQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQK FGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVH TECCHGDLLECADDRADLAKYICENQDSISSKLKECCE KPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKN YAEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYET TLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCE

		LFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLG KVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTP VSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAET FTFHADICTLSEKERQIKKQTALVELVKHKPKATKEQL KAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQ AALGLGGSGGSGGSGGSCSGPLGIEGGIISNQQITASSTH RALFGLQKWYPYYARLNKKGLINAWTAAENDRWPWI QINLQRKMRVTGVITQGAKRIGSPEYIKSYKIAYSNDGK TWAMYKVKGTNEDMVFRGNIDNNTPYANSFTPPIKAQ YVRLYPQVCRRHCTLRMELLGCELSGCSEPLGMKSGHI QDYQITASSIFRTLNMDMFTWEPRKARLDKQGKVNAW TSGHNDQSQWLQVDLLVPTKVTGIITQGAKDFGHVQF VGSYKLAYSNDGEHWTVYQDEKQRKDKVFQGNFDND THRKNVIDPPIYARHIRILPWSWYGRITLRSELLGC
84	miền giống EGF EDIL3 1[EDIL3]-HSA-C1-C2[EDIL3]	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTGSDAHKSEVAHRFKDLGEENFKALVLI AFAQYLQQSPFEDHVKLVNEVTEFAKTCVADESAENC DKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERN ECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLK KYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADK AACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFK AWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGD LLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSH CIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDV FLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCA AADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKC CKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVT KCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDD FAAFVEKCCKADDKETCFAEEGKKLVAASQAALGLGG

		SGGSGGSGGSCSGPLGIEGGIISNQQITASSTHRALFGLQ KWYPYYARLNKKGLINAWTAAENDRWPWIQINLQRK MRVTGVITQGAKRIGSPEYIKSYKIAYSNDGKTWAMY KVKGTNEDMVFRGNIDNNTPYANSFTPPIKAQYVRLYP QVCRRHCTLRMELLGCELSGCSEPLGMKSGHIQDYQIT ASSIFRTLNMDMFTWEPRKARLDKQGKVNAWTSGHN DQSQWLQVDLLVPTKVTGIITQGAKDFGHVQFVGSYK LAYSNDGEHWTVYQDEKQRKDKVFQGNFDNDTHRKN VIDPPIYARHIRILPWSWYGRITLRSELLGC
85	miền giống EGF EDIL3 2[EDIL3]-HSA-C1-C2[EDIL3]	SAGPCTPNPCHNGGTCEISEAYRGDTFIGYVCKCPRGFN GIHCQHGSDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVAASQAALGLGGSGGSGGS GGSCSGPLGIEGGIISNQQITASSTHRALFGLQKWYPYY ARLNKKGLINAWTAAENDRWPWIQINLQRKMRVTGVI TQGAKRIGSPEYIKSYKIAYSNDGKTWAMYKVKGTNE DMVFRGNIDNNTPYANSFTPPIKAQYVRLYPQVCRRHC TLRMELLGCELSGCSEPLGMKSGHIQDYQITASSIFRTL

		NMDMFTWEPRKARLDKQGKVNAWTSGHNDQSQWLQ VDLLVPTKVTGIITQGAKDFGHVQFVGSYKLAYSNDGE HWTVYQDEKQRKDKVFQGNFDNDTHRKNVIDPPIYAR HIRILPWSWYGRITLRSELLGC
86	miền giống EGF EDIL3 3[EDIL3]-HSA-C1-C2[EDIL3]	NINECEVEPCKNGGICTDLVANYSCECPGEFMGRNCQY KGSDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQSP FEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGD KLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDN PNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHP YFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDEL RDEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQR FPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRAD LAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEM PADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYA RRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAK VFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVR YTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMP CAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRR PCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIK KQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCK ADDKETCFAEEGKKLVAASQAALGLGGSGGSGGSGGS CSGPLGIEGGIISNQQITASSTHRALFGLQKWYPYYARL NKKGLINAWTAAENDRWPWIQINLQRKMRVTGVITQG AKRIGSPEYIKSYKIAYSNDGKTWAMYKVKGTNEDMV FRGNIDNNTPYANSFTPPIKAQYVRLYPQVCRRHCTLR MELLGCELSGCSEPLGMKSGHIQDYQITASSIFRTLNMD MFTWEPRKARLDKQGKVNAWTSGHNDQSQWLQVDL LVPTKVTGIITQGAKDFGHVQFVGSYKLAYSNDGEHW TVYQDEKQRKDKVFQGNFDNDTHRKNVIDPPIYARHIR ILPWSWYGRITLRSELLGC

87	miền giống EGF EDIL3 1-2[EDIL3]-HSA-C1-C2[EDIL3]	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHGSDAHKSEVAHRFKDLGEEN FKALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVAD ESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAK QEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDN EETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECC QAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHT ECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETT LEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCEL FEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK VGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPV SDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKA VMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAA LGLGGSGGSGGSGGSCSGPLGIEGGIISNQQITASSTHRA LFGLQKWYPYYARLNKKGLINAWTAAENDRWPWIQIN LQRKMRVTGVITQGAKRIGSPEYIKSYKIAYSNDGKTW AMYKVKGTNEDMVFRGNIDNNTPYANSFTPPIKAQYV RLYPQVCRRHCTLRMELLGCELSGCSEPLGMKSGHIQD YQITASSIFRTLNMDMFTWEPRKARLDKQGKVNAWTS GHNDQSQWLQVDLLVPTKVTGIITQGAKDFGHVQFVG SYKLAYSNDGEHWTVYQDEKQRKDKVFQGNFDNDTH RKNVIDPPIYARHIRILPWSWYGRITLRSELLGC
88	miền giống EGF EDIL3 2-3[EDIL3]-	SAGPCTPNPCHNGGTCEISEAYRGDTFIGYVCKCPRGFN GIHCQHNINECEVEPCKNGGICTDLVANYSCECPGEFM GRNCQYKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS

	HSA-C1-C2[EDIL3]	LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGLGGSGG SGGSGGSCSGPLGIEGGIISNQQITASSTHRALFGLQKW YPYYARLNKKGLINAWTAAENDRWPWIQINLQRKMR VTGVITQGAKRIGSPEYIKSYKIAYSNDGKTWAMYKVK GTNEDMVFRGNIDNNTPYANSFTPPIKAQYVRLYPQVC RRHCTLRMELLGCELSGCSEPLGMKSGHIQDYQITASSI FRTLNMDMFTWEPRKARLDKQGKVNAWTSGHNDQSQ WLQVDLLVPTKVTGIITQGAKDFGHVQFVGSYKLAYS NDGEHWTVYQDEKQRKDKVFQGNFDNDTHRKNVIDP PIYARHIRILPWSWYGRITLRSELLGC
89	miền giống EGF EDIL3 1-3[EDIL3]-HSA-C1-C2[EDIL3]	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTNINECEVEPCKNGGICTDLVANYSCECP GEFMGRNCQYKGSDAHKSEVAHRFKDLGEENFKALVL IAFAQYLQQSPFEDHVKLVNEVTEFAKTCVADESAENC DKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERN ECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLK KYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADK AACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFK

		AWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGD LLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSH CIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDV FLGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCA AADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGE YKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKC CKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVT KCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADI CTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDD FAAFVEKCCKADDKETCFAEEGKKLVAASQAALGLGG SGGSGGSGGSCSGPLGIEGGIISNQQITASSTHRALFGLQ KWYPYYARLNKKGLINAWTAAENDRWPWIQINLQRK MRVTGVITQGAKRIGSPEYIKSYKIAYSNDGKTWAMY KVKGTNEDMVFRGNIDNNTPYANSFTPPIKAQYVRLYP QVCRRHCTLRMELLGCELSGCSEPLGMKSGHIQDYQIT ASSIFRTLNMDMFTWEPRKARLDKQGKVNAWTSGHN DQSQWLQVDLLVPTKVTGIITQGAKDFGHVQFVGSYK LAYSNDGEHWTVYQDEKQRKDKVFQGNFDNDTHRKN VIDPPIYARHIRILPWSWYGRITLRSELLGC
13	FP050 Axit nucleic	gacatctgcgaccccaatccttgcgagaatggcggcatttgtctgcctggactggccga tggcagcttctcttgtgaatgccccgatggcttcacagaccccaattgcagctctgtggtg gaagtggccagcgacgaggaagaacctacaagcgctggcccctgcacacccaatcc atgtcataatggcggaacctgcgagatcagcgaggcctacagaggcgataccttcatc ggctacgtgtgcaagtgccccagaggcttcaatggcatccactgccagcacaacatca acgagtgcgaggtggaaccatgcaagaacggcggcatctgtaccgacctggtggcca attactcttgcgagtgccctggcgagttcatgggcagaaactgccagtacaagggatcc gacgctcacaagtctgaggtggcccacagattcaaggacctgggcgaagagaacttca aggccctggtgctgatcgccttcgctcagtatctgcagcagagccctttcgaggaccac gtgaagctggtcaacgaagtgaccgagttcgccaagacctgtgtggccgatgagagc gccgagaactgtgacaagagcctgcacacactgttcggcgacaagctgtgtaccgtgg ccacactgagagaaacctacggcgagatggccgactgctgtgccaagcaagagccc

		gagagaaacgagtgcttcctgcagcacaaggacgacaaccccaacctgcctagactc gtgcgacccgaagtggatgtgatgtgcaccgcctttcacgacaacgaggaaaccttcct gaagaagtacctgtacgagatcgccagacggcacccctacttttatgcccctgagctgc tgttcttcgccaagcggtataaggccgccttcaccgaatgttgccaggccgctgataag gctgcctgtctgctgcctaagctggacgagctgagagatgagggcaaagccagctctg ccaagcagagactgaagtgcgccagcctgcagaagttcggcgagagagcttttaagg cctgggccgttgccagactgagccagagatttcctaaggccgagtttgccgaggtgtcc aagctcgtgaccgatctgacaaaggtgcacaccgagtgctgtcacggcgatctgctgg aatgtgccgacgatagagccgacctggccaagtatatctgcgagaaccaggacagcat cagcagcaagctgaaagagtgctgcgagaagcccctgctggaaaagtctcactgtatc gccgaggtcgagaacgacgagatgcctgctgatctgcctagcctggccgccgatttcg tggaaagcaaggatgtgtgcaagaactacgccgaggccaaagatgtgtttctgggcat gtttctgtatgagtacgcccgcagacaccccgactattctgtggttctgctgctgcggctg gccaagacatacgagacaaccctggaaaaatgctgcgccgctgccgatcctcacgagt gttatgccaaggtgttcgacgagttcaagccactggtggaagaaccccagaacctgatc aagcagaactgcgagctgttcgagcagctgggcgagtacaagttccagaatgccctgc tcgtgcggtacaccaagaaagtgcctcaggtgtccacacctacactggttgaggtgtcc cggaatctgggcaaagtgggcagcaagtgttgcaagcaccctgaggccaagagaatg ccttgcgccgaggattacctgagcgtggtgctgaatcagctgtgcgtgctgcacgagaa aacccctgtgtccgacagagtgaccaagtgctgtaccgagagcctcgtgaacagaagg ccttgctttagcgccctggaagtggacgagacatacgtgcccaaagagttcaacgccg agacattcaccttccacgccgatatctgcaccctgtccgagaaagagcggcagatcaa gaagcagacagccctggtcgagctggttaagcacaagcccaaggccaccaaagaac agctgaaggccgtgatggacgacttcgccgcctttgtcgagaagtgctgcaaggccga cgacaaagagacatgcttcgccgaagagggcaagaaactggtggctgcctctcaggc tgctctcggacttggaggaagcggaggatctggcggttccggaggaagttgttctggcc ctcttggcatcgaaggcggcatcatcagcaatcagcagatcaccgccagcagcaccca cagagcactgtttggactgcagaaatggtatccctactacgcccggctgaacaagaag ggcctgattaacgcctggacagccgccgagaatgacagatggccctggattcagatca acctgcagcggaagatgagagtgaccggcgttatcacacagggcgccaaaagaatcg gcagccccgagtacatcaagagctacaagatcgcctacagcaacgacggcaagacct gggccatgtacaaagtgaagggcaccaacgaggacatggtgttccggggcaacatcg

		acaacaacacccttacgccaacagcttcacccctcctatcaaggcccagtacgtgcgg ctgtaccctcaagtgtgcagaaggcactgtaccctgagaatggaactgctgggctgcga actgtctggctgttctgagccactgggcatgaagtccggccacatccaggattaccagat cacagcctccagcatcttcagaaccctgaacatggatatgttcacctgggagccccgga aggccagactggataagcagggaaaagtgaatgcctggaccagcggccacaacgac cagtctcaatggctgcaagtggacctgctggtgcccaccaaagtgaccggaatcattac tcagggcgcaaaggacttcggccacgtgcagtttgtgggctcctacaagctggcctact ccaacgatggcgagcactggacagtgtaccaggacgagaagcagcgcaaggataag gtgttccagggaaacttcgataacgatacccaccggaagaacgtgatcgaccctccaat ctacgccagacacatcagaatcctgccttggtcttggtacggcagaatcaccctgagatc cgagctgctgggatgc
90	Axit nucleic có Seq ID NO: 84	gacatctgcgaccccaaccctgcgagaacggcggcatctgcctgcccggcctggcc gacggcagcttcagctgcgagtgccccgacggcttcaccgaccccaactgcagcagc gtggtggaggtggccagcgacgaggaggagcccaccggcagcgacgcccacaaga gcgaggtggcccaccggttcaaggacctgggcgaggagaacttcaaggccctggtgc tgatcgccttcgcccagtacctgcagcagagccccttcgaggaccacgtgaagctggt gaacgaggtgaccgagttcgccaagacctgcgtggccgacgagagcgccgagaact gcgacaagagcctgcacaccctgttcggcgacaagctgtgcaccgtggccaccctgc gggagacctacggcgagatggccgactgctgcgccaagcaggagcccgagcggaa cgagtgcttcctgcagcacaaggacgacaaccccaacctgccccggctggtgcggcc cgaggtggacgtgatgtgcaccgccttccacgacaacgaggagaccttcctgaagaa gtacctgtacgagatcgcccggcggcacccctacttctacgcccccgagctgctgttctt cgccaagcggtacaaggccgccttcaccgagtgctgccaggccgccgacaaggccg cctgcctgctgcccaagctggacgagctgcgggacgagggcaaggccagcagcgcc aagcagcggctgaagtgcgccagcctgcagaagttcggcgagcgggccttcaaggc ctgggccgtggcccggctgagccagcggttccccaaggccgagttcgccgaggtga gcaagctggtgaccgacctgaccaaggtgcacaccgagtgctgccacggcgacctgc tggagtgcgccgacgaccgggccgacctggccaagtacatctgcgagaaccaggac agcatcagcagcaagctgaaggagtgctgcgagaagcccctgctggagaagagcca ctgcatcgccgaggtggagaacgacgagatgcccgccgacctgcccagcctggccg ccgacttcgtggagagcaaggacgtgtgcaagaactacgccgaggccaaggacgtgt

		tcctgggcatgttcctgtacgagtacgcccggcggcaccccgactacagcgtggtgct gctgctgcggctggccaagacctacgagaccaccctggagaagtgctgcgccgccgc cgacccccacgagtgctacgccaaggtgttcgacgagttcaagcccctggtggagga gccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagtacaa gttccagaacgccctgctggtgcggtacaccaagaaggtgccccaggtgagcacccc caccctggtggaggtgagccggaacctgggcaaggtgggcagcaagtgctgcaagc accccgaggccaagcggatgccctgcgccgaggactacctgagcgtggtgctgaacc agctgtgcgtgctgcacgagaagacccccgtgagcgaccgggtgaccaagtgctgca ccgagagcctggtgaaccggcggccctgcttcagcgccctggaggtggacgagacct acgtgcccaaggagttcaacgccgagaccttcaccttccacgccgacatctgcaccctg agcgagaaggagcggcagatcaagaagcagaccgccctggtggagctggtgaagc acaagcccaaggccaccaaggagcagctgaaggccgtgatggacgacttcgccgcc ttcgtggagaagtgctgcaaggccgacgacaaggagacctgcttcgccgaggagggc aagaagctggtggccgccagccaggccgccctgggcctgggcggcagcggcggca gcggcggcagcggcggcagctgcagcggccccctgggcatcgagggcggcatcat cagcaaccagcagatcaccgccagcagcacccaccgggccctgttcggcctgcaga agtggtaccctactacgcccggctgaacaagaagggcctgatcaacgcctggaccg ccgccgagaacgaccggtggccctggatccagatcaacctgcagcggaagatgcgg gtgaccggcgtgatcacccagggcgccaagcggatcggcagccccgagtacatcaa gagctacaagatcgcctacagcaacgacggcaagacctgggccatgtacaaggtgaa gggcaccaacgaggacatggtgttccggggcaacatcgacaacaacacccccctacgc caacagcttcacccccccccatcaaggcccagtacgtgcggctgtacccccaggtgtgc cggcggcactgcaccctgcggatggagctgctgggctgcgagctgagcggctgcag cgagcccctgggcatgaagagcggccacatccaggactaccagatcaccgccagca gcatcttccggaccctgaacatggacatgttcacctgggagccccggaaggcccggct ggacaagcagggcaaggtgaacgcctggaccagcggccacaacgaccagagccag tggctgcaggtggacctgctggtgcccaccaaggtgaccggcatcatcacccagggc gccaaggacttcggccacgtgcagttcgtgggcagctacaagctggcctacagcaac gacggcgagcactggaccgtgtaccaggacgagaagcagcggaaggacaaggtgtt ccagggcaacttcgacaacgacacccaccggaagaacgtgatcgacccccccatcta cgcccggcacatccggatcctgccctggagctggtacggccggatcaccctgcggag cgagctgctgggctgc

91	Axit nucleic có Seq ID NO: 85	agcgccggcccctgcacccccaaccccctgccacaacggcggcacctgcgagatcag cgaggcctaccggggcgacaccttcatcggctacgtgtgcaagtgcccccggggcttc aacggcatccactgccagcacggcagcgacgcccacaagagcgaggtggcccacc ggttcaaggacctgggcgaggagaacttcaaggccctggtgctgatcgccttcgccca gtacctgcagcagagcccttcgaggaccacgtgaagctggtgaacgaggtgaccga gttcgccaagacctgcgtggccgacgagagcgccgagaactgcgacaagagcctgc acaccctgttcggcgacaagctgtgcaccgtggccaccctgcgggagacctacggcg agatggccgactgctgcgccaagcaggagcccgagcggaacgagtgcttcctgcag cacaaggacgacaaccccaacctgccccggctggtgcggcccgaggtggacgtgat gtgcaccgccttccacgacaacgaggagaccttcctgaagaagtacctgtacgagatc gcccggcggcacccctacttctacgccccgagctgctgttcttcgccaagcggtacaa ggccgccttcaccgagtgctgccaggccgccgacaaggccgcctgcctgctgcccaa gctggacgagctgcgggacgagggcaaggccagcagcgccaagcagcggctgaa gtgcgccagcctgcagaagttcggcgagcgggccttcaaggcctgggccgtggccc ggctgagccagcggttccccaaggccgagttcgccgaggtgagcaagctggtgaccg acctgaccaaggtgcacaccgagtgctgccacggcgacctgctggagtgcgccgacg accgggccgacctggccaagtacatctgcgagaaccaggacagcatcagcagcaag ctgaaggagtgctgcgagaagcccctgctggagaagagccactgcatcgccgaggtg gagaacgacgagatgcccgccgacctgcccagcctggccgccgacttcgtggagag caaggacgtgtgcaagaactacgccgaggccaaggacgtgttcctgggcatgttcctg tacgagtacgcccggcggcaccccgactacagcgtggtgctgctgctgcggctggcc aagacctacgagaccaccctggagaagtgctgcgccgccgccgacccccacgagtg ctacgccaaggtgttcgacgagttcaagcccctggtggaggagccccagaacctgatc aagcagaactgcgagctgttcgagcagctgggcgagtacaagttccagaacgccctg ctggtgcggtacaccaagaaggtgccccaggtgagcacccccacccctggtggaggtg agccggaacctgggcaaggtgggcagcaagtgctgcaagcaccccgaggccaagc ggatgccctgcgccgaggactacctgagcgtggtgctgaaccagctgtgcgtgctgca cgagaagacccccgtgagcgaccgggtgaccaagtgctgcaccgagagcctggtga accggcggccctgcttcagcgccctggaggtggacgagacctacgtgcccaaggagt tcaacgccgagaccttcaccttccacgccgacatctgcaccctgagcgagaaggagcg gcagatcaagaagcagaccgccctggtggagctggtgaagcacaagcccaaggcca ccaaggagcagctgaaggccgtgatggacgacttcgccgccttcgtggagaagtgct

		gcaaggccgacgacaaggagacctgcttcgccgaggagggcaagaagctggtggc cgccagccaggccgccctgggcctgggcggcagcggcggcagcggcggcagcgg cggcagctgcagcggcccccctgggcatcgagggcggcatcatcagcaaccagcaga tcaccgccagcagcacccaccgggccctgttcggcctgcagaagtggtacccctacta cgcccggctgaacaagaagggcctgatcaacgcctggaccgccgccgagaacgacc ggtggccctggatccagatcaacctgcagcggaagatgcgggtgaccggcgtgatca cccagggcgccaagcggatcggcagccccgagtacatcaagagctacaagatcgcct acagcaacgacggcaagacctgggccatgtacaaggtgaagggcaccaacgaggac atggtgttccggggcaacatcgacaacaacacccccctacgccaacagcttcacccccc ccatcaaggcccagtacgtgcggctgtacccccaggtgtgccggcggcactgcaccct gcggatggagctgctgggctgcgagctgagcggctgcagcgagcccctgggcatga agagcggccacatccaggactaccagatcaccgccagcagcatcttccggaccctga acatggacatgttcacctgggagccccggaaggcccggctggacaagcagggcaag gtgaacgcctggaccagcggccacaacgaccagagccagtggctgcaggtggacct gctggtgcccaccaaggtgaccggcatcatcacccagggcgccaaggacttcggcca cgtgcagttcgtgggcagctacaagctggcctacagcaacgacggcgagcactggac cgtgtaccaggacgagaagcagcggaaggacaaggtgttccagggcaacttcgacaa cgacacccaccggaagaacgtgatcgacccccccatctacgcccggcacatccggat cctgccctggagctggtacggccggatcaccctgcggagcgagctgctgggctgc
92	Axit nucleic có Seq ID NO: 86	aacatcaacgagtgcgaggtggagccctgcaagaacggcggcatctgcaccgacctg gtggccaactacagctgcgagtgccccggcgagttcatgggccggaactgccagtac aagggcagcgacgcccacaagagcgaggtggcccaccggttcaaggacctgggcg aggagaacttcaaggccctggtgctgatcgccttcgcccagtacctgcagcagagccc cttcgaggaccacgtgaagctggtgaacgaggtgaccgagttcgccaagacctgcgtg gccgacgagagcgccgagaactgcgacaagagcctgcacaccctgttcggcgacaa gctgtgcaccgtggccaccctgcgggagacctacggcgagatggccgactgctgcgc caagcaggagcccgagcggaacgagtgcttcctgcagcacaaggacgacaacccca acctgccccggctggtgcggcccgaggtggacgtgatgtgcaccgccttccacgaca acgaggagaccttcctgaagaagtacctgtacgagatcgcccggcggcacccctactt ctacgcccccgagctgctgttcttcgccaagcggtacaaggccgccttcaccgagtgct gccaggccgccgacaaggccgcctgcctgctgcccaagctggacgagctgcgggac

		gagggcaaggccagcagcgccaagcagcggctgaagtgcgccagcctgcagaagtt cggcgagcgggccttcaaggcctgggccgtggcccggctgagccagcggttcccca aggccgagttcgccgaggtgagcaagctggtgaccgacctgaccaaggtgcacacc gagtgctgccacggcgacctgctggagtgcgccgacgaccgggccgacctggccaa gtacatctgcgagaaccaggacagcatcagcagcaagctgaaggagtgctgcgagaa gcccctgctggagaagagccactgcatcgccgaggtggagaacgacgagatgcccg ccgacctgcccagcctggccgccgacttcgtggagagcaaggacgtgtgcaagaact acgccgaggccaaggacgtgttcctgggcatgttcctgtacgagtacgcccggcggca ccccgactacagcgtggtgctgctgctgcggctggccaagacctacgagaccaccctg gagaagtgctgcgccgccgccgacccccacgagtgctacgccaaggtgttcgacgag ttcaagcccctggtggaggagccccagaacctgatcaagcagaactgcgagctgttcg agcagctgggcgagtacaagttccagaacgccctgctggtgcggtacaccaagaagg tgccccaggtgagcacccccaccctggtggaggtgagccggaacctgggcaaggtg ggcagcaagtgctgcaagcaccccgaggccaagcggatgccctgcgccgaggacta cctgagcgtggtgctgaaccagctgtgcgtgctgcacgagaagacccccgtgagcga ccgggtgaccaagtgctgcaccgagagcctggtgaaccggcggccctgcttcagcgc cctggaggtggacgagacctacgtgcccaaggagttcaacgccgagaccttcaccttc cacgccgacatctgcaccctgagcgagaaggagcggcagatcaagaagcagaccgc cctggtggagctggtgaagcacaagcccaaggccaccaaggagcagctgaaggccg tgatggacgacttcgccgccttcgtggagaagtgctgcaaggccgacgacaaggaga cctgcttcgccgaggagggcaagaagctggtggccgccagccaggccgccctgggc ctgggcggcagcggcggcagcggcggcagcggcggcagctgcagcggcccctg ggcatcgagggcggcatcatcagcaaccagcagatcaccgccagcagcacccaccg ggccctgttcggcctgcagaagtggtaccctactacgcccggctgaacaagaaggg cctgatcaacgcctggaccgccgccgagaacgaccggtggccctggatccagatcaa cctgcagcggaagatgcgggtgaccggcgtgatcacccagggcgccaagcggatcg gcagccccgagtacatcaagagctacaagatcgcctacagcaacgacggcaagacct gggccatgtacaaggtgaagggcaccaacgaggacatggtgttccggggcaacatcg acaacaacacccctacgccaacagcttcacccccccccatcaaggcccagtacgtgcg gctgtaccccaggtgtgccggcggcactgcaccctgcggatggagctgctgggctg cgagctgagcggctgcagcgagcccctgggcatgaagagcggccacatccaggact accagatcaccgccagcagcatcttccggaccctgaacatggacatgttcacctggga

		gccccggaaggcccggctggacaagcagggcaaggtgaacgcctggaccagcggc cacaacgaccagagccagtggctgcaggtggacctgctggtgcccaccaaggtgacc ggcatcatcacccagggcgccaaggacttcggccacgtgcagttcgtgggcagctac aagctggcctacagcaacgacggcgagcactggaccgtgtaccaggacgagaagca gcggaaggacaaggtgttccagggcaacttcgacaacgacacccaccggaagaacg tgatcgacccccccatctacgcccggcacatccggatcctgccctggagctggtacgg ccggatcaccctgcggagcgagctgctgggctgc
93	Axit nucleic có Seq ID NO: 87	gacatctgcgaccccaacccctgcgagaacggcggcatctgcctgcccggcctggcc gacggcagcttcagctgcgagtgccccgacggcttcaccgaccccaactgcagcagc gtggtggaggtggccagcgacgaggaggagcccaccagcgccggcccctgcaccc ccaaccctgccacaacggcggcacctgcgagatcagcgaggcctaccggggcgac accttcatcggctacgtgtgcaagtgcccccggggcttcaacggcatccactgccagca cggcagcgacgcccacaagagcgaggtggcccaccggttcaaggacctgggcgag gagaacttcaaggccctggtgctgatcgccttcgcccagtacctgcagcagagcccctt cgaggaccacgtgaagctggtgaacgaggtgaccgagttcgccaagacctgcgtggc cgacgagagcgccgagaactgcgacaagagcctgcacaccctgttcggcgacaagc tgtgcaccgtggccaccctgcgggagacctacggcgagatggccgactgctgcgcca agcaggagcccgagcggaacgagtgcttcctgcagcacaaggacgacaaccccaac ctgccccggctggtgcggcccgaggtggacgtgatgtgcaccgccttccacgacaac gaggagaccttcctgaagaagtacctgtacgagatcgcccggcggcacccctacttcta cgcccccgagctgctgttcttcgccaagcggtacaaggccgccttcaccgagtgctgc caggccgccgacaaggccgcctgcctgctgcccaagctggacgagctgcgggacga gggcaaggccagcagcgccaagcagcggctgaagtgcgccagcctgcagaagttcg gcgagcgggccttcaaggcctgggccgtggcccggctgagccagcggttccccaag gccgagttcgccgaggtgagcaagctggtgaccgacctgaccaaggtgcacaccga gtgctgccacggcgacctgctggagtgcgccgacgaccgggccgacctggccaagt acatctgcgagaaccaggacagcatcagcagcaagctgaaggagtgctgcgagaag cccctgctggagaagagccactgcatcgccgaggtggagaacgacgagatgcccgc cgacctgcccagcctggccgccgacttcgtggagagcaaggacgtgtgcaagaacta cgccgaggccaaggacgtgttcctgggcatgttcctgtacgagtacgcccggcggcac cccgactacagcgtggtgctgctgctgcggctggccaagacctacgagaccaccctg

		gagaagtgctgcgccgccgccgacccccacgagtgctacgccaaggtgttcgacgag ttcaagcccctggtggaggagccccagaacctgatcaagcagaactgcgagctgttcg agcagctgggcgagtacaagttccagaacgccctgctggtgcggtacaccaagaagg tgccccaggtgagcacccccaccctggtggaggtgagccggaacctgggcaaggtg ggcagcaagtgctgcaagcaccccgaggccaagcggatgccctgcgccgaggacta cctgagcgtggtgctgaaccagctgtgcgtgctgcacgagaagacccccgtgagcga ccgggtgaccaagtgctgcaccgagagcctggtgaaccggcggccctgcttcagcgc cctggaggtggacgagacctacgtgcccaaggagttcaacgccgagaccttcaccttc cacgccgacatctgcaccctgagcgagaaggagcggcagatcaagaagcagaccgc cctggtggagctggtgaagcacaagcccaaggccaccaaggagcagctgaaggccg tgatggacgacttcgccgccttcgtggagaagtgctgcaaggccgacgacaaggaga cctgcttcgccgaggagggcaagaagctggtggccgccagccaggccgccctgggc ctgggcggcagcggcggcagcggcggcagcggcggcagctgcagcggcccctg ggcatcgagggcggcatcatcagcaaccagcagatcaccgccagcagcacccaccg ggccctgttcggcctgcagaagtggtacccctactacgcccggctgaacaagaaggg cctgatcaacgcctggaccgccgccgagaacgaccggtggccctggatccagatcaa cctgcagcggaagatgcgggtgaccggcgtgatcacccagggcgccaagcggatcg gcagccccgagtacatcaagagctacaagatcgcctacagcaacgacggcaagacct gggccatgtacaaggtgaagggcaccaacgaggacatggtgttccggggcaacatcg acaacaacacccctacgccaacagcttcacccccccccatcaaggcccagtacgtgcg gctgtaccccaggtgtgccggcggcactgcaccctgcggatggagctgctgggctg cgagctgagcggctgcagcgagcccctgggcatgaagagcggccacatccaggact accagatcaccgccagcagcatcttccggaccctgaacatggacatgttcacctggga gccccggaaggcccggctggacaagcagggcaaggtgaacgcctggaccagcggc cacaacgaccagagccagtggctgcaggtggacctgctggtgcccaccaaggtgacc ggcatcatcacccagggcgccaaggacttcggccacgtgcagttcgtgggcagctac aagctggcctacagcaacgacggcgagcactggaccgtgtaccaggacgagaagca gcggaaggacaaggtgttccagggcaacttcgacaacgacacccaccggaagaacg tgatcgacccccccatctacgcccggcacatccggatcctgccctggagctggtacgg ccggatcaccctgcggagcgagctgctgggctgc

94	Axit nucleic có Seq ID NO: 88	agcgccggcccctgcacccccaaccccctgccacaacggcggcacctgcgagatcag cgaggcctaccggggcgacaccttcatcggctacgtgtgcaagtgcccccggggcttc aacggcatccactgccagcacaacatcaacgagtgcgaggtggagccctgcaagaac ggcggcatctgcaccgacctggtggccaactacagctgcgagtgccccggcgagttc atgggccggaactgccagtacaagggcagcgacgcccacaagagcgaggtggccc accggttcaaggacctgggcgaggagaacttcaaggccctggtgctgatcgccttcgc ccagtacctgcagcagagccccttcgaggaccacgtgaagctggtgaacgaggtgac cgagttcgccaagacctgcgtggccgacgagagcgccgagaactgcgacaagagcc tgcacaccctgttcggcgacaagctgtgcaccgtggccaccctgcgggagacctacg gcgagatggccgactgctgcgccaagcaggagcccgagcggaacgagtgcttcctg cagcacaaggacgacaaccccaacctgccccggctggtgcggcccgaggtggacgt gatgtgcaccgccttccacgacaacgaggagaccttcctgaagaagtacctgtacgag atcgcccggcggcacccctacttctacgccccgagctgctgttcttcgccaagcggta caaggccgccttcaccgagtgctgccaggccgccgacaaggccgcctgcctgctgcc caagctggacgagctgcgggacgagggcaaggccagcagcgccaagcagcggctg aagtgcgccagcctgcagaagttcggcgagcgggccttcaaggcctgggccgtggcc cggctgagccagcggttccccaaggccgagttcgccgaggtgagcaagctggtgacc gacctgaccaaggtgcacaccgagtgctgccacggcgacctgctggagtgcgccgac gaccgggccgacctggccaagtacatctgcgagaaccaggacagcatcagcagcaa gctgaaggagtgctgcgagaagcccctgctggagaagagccactgcatcgccgaggt ggagaacgacgagatgcccgccgacctgcccagcctggccgccgacttcgtggaga gcaaggacgtgtgcaagaactacgccgaggccaaggacgtgttcctgggcatgttcct gtacgagtacgcccggcggcaccccgactacagcgtggtgctgctgcggctggc caagacctacgagaccaccctggagaagtgctgcgccgccgccgaccccacgagt gctacgccaaggtgttcgacgagttcaagcccctggtggaggagccccagaacctgat caagcagaactgcgagctgttcgagcagctgggcgagtacaagttccagaacgccct gctggtgcggtacaccaagaaggtgccccaggtgagcacccccaccctggtggaggt gagccggaacctgggcaaggtgggcagcaagtgctgcaagcaccccgaggccaag cggatgccctgcgccgaggactacctgagcgtggtgctgaaccagctgtgcgtgctgc acgagaagacccccgtgagcgaccgggtgaccaagtgctgcaccgagagcctggtg aaccggcggccctgcttcagcgccctggaggtggacgagacctacgtgcccaagga gttcaacgccgagaccttcaccttccacgccgacatctgcaccctgagcgagaaggag

		cggcagatcaagaagcagaccgccctggtggagctggtgaagcacaagcccaaggc caccaaggagcagctgaaggccgtgatggacgacttcgccgccttcgtggagaagtg ctgcaaggccgacgacaaggagacctgcttcgccgaggagggcaagaagctggtgg ccgccagccaggccgccctgggcctgggcggcagcggcggcagcggcggcagcg gcggcagctgcagcggcccctgggcatcgagggcggcatcatcagcaaccagcag atcaccgccagcagcacccaccgggccctgttcggcctgcagaagtggtacccctact acgcccggctgaacaagaagggcctgatcaacgcctggaccgccgccgagaacgac cggtggccctggatccagatcaacctgcagcggaagatgcgggtgaccggcgtgatc acccagggcgccaagcggatcggcagccccgagtacatcaagagctacaagatcgc ctacagcaacgacggcaagacctgggccatgtacaaggtgaagggcaccaacgagg acatggtgttccggggcaacatcgacaacaacacccctacgccaacagcttcacccc ccccatcaaggcccagtacgtgcggctgtacccccaggtgtgccggcggcactgcac cctgcggatggagctgctgggctgcgagctgagcggctgcagcgagcccctgggcat gaagagcggccacatccaggactaccagatcaccgccagcagcatcttccggaccct gaacatggacatgttcacctgggagccccggaaggcccggctggacaagcagggca aggtgaacgcctggaccagcggccacaacgaccagagccagtggctgcaggtggac ctgctggtgcccaccaaggtgaccggcatcatcacccagggcgccaaggacttcggc cacgtgcagttcgtgggcagctacaagctggcctacagcaacgacggcgagcactgg accgtgtaccaggacgagaagcagcggaaggacaaggtgttccagggcaacttcgac aacgacacccaccggaagaacgtgatcgacccccccatctacgcccggcacatccgg atcctgccctggagctggtacggccggatcaccctgcggagcgagctgctgggctgc
95	Ácit nucleic có Seq ID NO: 89	gacatctgcgaccccaaccccctgcgagaacggcggcatctgcctgcccggcctggcc gacggcagcttcagctgcgagtgccccgacggcttcaccgaccccaactgcagcagc gtggtggaggtggccagcgacgaggaggagcccaccaacatcaacgagtgcgaggt ggagccctgcaagaacggcggcatctgcaccgacctggtggccaactacagctgcga gtgccccggcgagttcatgggccggaactgccagtacaagggcagcgacgcccaca agagcgaggtggcccaccggttcaaggacctgggcgaggagaacttcaaggccctg gtgctgatcgccttcgcccagtacctgcagcagagccccttcgaggaccacgtgaagc tggtgaacgaggtgaccgagttcgccaagacctgcgtggccgacgagagcgccgag aactgcgacaagagcctgcacaccctgttcggcgacaagctgtgcaccgtggccacc ctgcgggagacctacggcgagatggccgactgctgcgccaagcaggagcccgagc

		ggaacgagtgcttcctgcagcacaaggacgacaaccccaacctgccccggctggtgc ggcccgaggtggacgtgatgtgcaccgccttccacgacaacgaggagaccttcctga agaagtacctgtacgagatcgcccggcggcacccctacttctacgcccccgagctgct gttcttcgccaagcggtacaaggccgccttcaccgagtgctgccaggccgccgacaag gccgcctgcctgctgcccaagctggacgagctgcgggacgagggcaaggccagca gcgccaagcagcggctgaagtgcgccagcctgcagaagttcggcgagcgggccttc aaggcctgggccgtggcccggctgagccagcggttccccaaggccgagttcgccga ggtgagcaagctggtgaccgacctgaccaaggtgcacaccgagtgctgccacggcg acctgctggagtgcgccgacgaccgggccgacctggccaagtacatctgcgagaacc aggacagcatcagcagcaagctgaaggagtgctgcgagaagcccctgctggagaag agccactgcatcgccgaggtggagaacgacgagatgcccgccgacctgcccagcct ggccgccgacttcgtggagagcaaggacgtgtgcaagaactacgccgaggccaagg acgtgttcctgggcatgttcctgtacgagtacgcccggcggcaccccgactacagcgtg gtgctgctgctgcggctggccaagacctacgagaccaccctggagaagtgctgcgcc gccgccgacccccacgagtgctacgccaaggtgttcgacgagttcaagcccctggtg gaggagccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcga gtacaagttccagaacgccctgctggtgcggtacaccaagaaggtgccccaggtgag cacccccaccctggtggaggtgagccggaacctgggcaaggtgggcagcaagtgct gcaagcaccccgaggccaagcggatgccctgcgccgaggactacctgagcgtggtg ctgaaccagctgtgcgtgctgcacgagaagacccccgtgagcgaccgggtgaccaag tgctgcaccgagagcctggtgaaccggcggccctgcttcagcgccctggaggtggac gagacctacgtgcccaaggagttcaacgccgagaccttcaccttccacgccgacatct gcaccctgagcgagaaggagcggcagatcaagaagcagaccgccctggtggagctg gtgaagcacaagcccaaggccaccaaggagcagctgaaggccgtgatggacgactt cgccgccttcgtggagaagtgctgcaaggccgacgacaaggagacctgcttcgccga ggagggcaagaagctggtggccgccagccaggccgccctgggcctgggcggcagc ggcggcagcggcggcagcggcggcagctgcagcggcccccctgggcatcgagggc ggcatcatcagcaaccagcagatcaccgccagcagcacccaccgggccctgttcggc ctgcagaagtggtaccctactacgcccggctgaacaagaagggcctgatcaacgcct ggaccgccgccgagaacgaccggtggccctggatccagatcaacctgcagcggaag atgcgggtgaccggcgtgatcacccagggcgccaagcggatcggcagccccgagta catcaagagctacaagatcgcctacagcaacgacggcaagacctgggccatgtacaa

		ggtgaagggcaccaacgaggacatggtgttccggggcaacatcgacaacaacaccc cctacgccaacagcttcacccccccccatcaaggcccagtacgtgcggctgtaccccca ggtgtgccggcggcactgcaccctgcggatggagctgctgggctgcgagctgagcg gctgcagcgagcccctgggcatgaagagcggccacatccaggactaccagatcacc gccagcagcatcttccggaccctgaacatggacatgttcacctgggagccccggaagg cccggctggacaagcagggcaaggtgaacgcctggaccagcggccacaacgacca gagccagtggctgcaggtggacctgctggtgcccaccaaggtgaccggcatcatcac ccagggcgccaaggacttcggccacgtgcagttcgtgggcagctacaagctggccta cagcaacgacggcgagcactggaccgtgtaccaggacgagaagcagcggaaggac aaggtgttccagggcaacttcgacaacgacacccaccggaagaacgtgatcgacccc cccatctacgcccggcacatccggatcctgccctggagctggtacggccggatcaccc tgcggagcgagctgctgggctgc
14	FP060 EGF-C1-C2-Fc [S354C, T366W]	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKCVEPLGMENGNIANSQIAASSVRVTFLGLQH WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCELNGCANPLGLKNNSIPDKQITA SSSYKTWGLHLFSWNPSYARLDKQGNFNAWVAGSYG NDQWLQVDLGSSKEVTGIITQGARNFGSVQFVASYKV AYSNDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNLF ETPILARYVRILPVAWHNRIALRLELLGCGGGGTDKTH TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVV DVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTY RVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISK AKGQPREPQVYTLPPCREEMTKNQVSLWCLVKGFYPS DIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVD KSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK
15	FP05C Axit nucleic	ctggacatctgcagcaagaacccctgccacaacggcggcctgtgcgaagagatcagc caggaagtgcggggcgacgtgttccccagctacacctgtacctgcctgaagggctacg ccggcaaccactgcgagactaagtgcgtggaacccctgggcatggaaaacggcaata

		ttgccaacagccagatcgccgccagctccgtgcgcgtgacctttctgggactgcagca ctgggtgcccgagctggccagactgaacagagccggcatggtgaacgcctggaccc ccagcagcaacgacgacaacccttggatccaggtgaacctgctgcggcggatgtggg tgacaggcgtggtgacacagggcgccagcagactggccagccacgagtacctgaag gcctttaaggtggcctacagcctgaacggccacgagttcgacttcatccacgacgtgaa caagaaacacaaagaatttgtgggcaactggaacaagaacgccgtgcacgtgaacct gttcgagacacccgtggaagcccagtacgtgcggctgtaccccaccagctgccacac cgcctgcaccctgagattcgagctgctgggctgcgagctgaacggctgcgccaaccc cctgggcctgaagaacaacagcatccccgacaagcagatcaccgcctccagcagcta caagacctggggcctgcacctgttcagctggaaccccagctacgcccggctggacaa gcagggcaacttcaacgcctgggtggccggcagctacggcaacgaccagtggctgc aggtggacctgggcagcagcaaagaagtgaccggcatcatcacccaggggccaga aacttcggcagcgtgcagttcgtggccagctacaaagtggcctactccaacgacagcg ccaactggaccgagtaccaggacccccggaccggcagctccaagatcttccccggca actgggacaaccacagccacaagaagaatctgttcgaaaccccccatcctggccagata cgtgcggatcctgcccgtggcctggcacaaccggatcgccctgagactggaactgctg ggatgtgggggaggcggtaccgacaagacccacacctgccccccctgcccagcccc agagctgctgggcggaccctccgtgttcctgttcccccccaagcccaaggacaccctg atgatcagcaggacccccgaggtgacctgcgtggtggtggacgtgagccacgaggac ccagaggtgaagttcaactggtacgtggacggcgtggaggtgcacaacgccaagacc aagcccagagaggagcagtacaacagcacctacagggtggtgtccgtgctgaccgtg ctgcaccaggactggctgaacggcaaggaatacaagtgcaaggtctccaacaaggcc ctgccagccccccatcgaaaagaccatcagcaaggccaagggccagccacgggagcc ccaggtgtacaccctgcccccctgccgggaggagatgaccaagaaccaggtgtccct gtggtgtctggtgaagggcttctaccccagcgacatcgccgtggagtgggagagcaac ggccagcccgagaacaactacaagaccacccccccagtgctggacagcgacggcag cttcttcctgtacagcaagctgaccgtggacaagtccaggtggcagcagggcaacgtgt tcagctgcagcgtgatgcacgaggccctgcacaaccactacacccagaagagcctga gcctgtcccccggcaag
16	FP070	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDKTHTCPPCPAPELLGGPSVFLFPPKPKDT

	EGF-Fc-C1-C2	LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA KTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS NKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQV SLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSD GSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQ KSLSLSPGKGGSGGSGGSGGSCVEPLGMENGNIANSQI AASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCAN PLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYARLD KQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGA RNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFP GNWDNHSHKKNLFETPILARYVRILPVAWHNRIALRLE LLGCGSHHHHHH
17	FP070 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgacaagacccacacctgccccccctgccca gccccagagctgctgggcggaccctccgtgttcctgttcccccccaagcccaaggaca ccctgatgatcagcaggacccccgaggtgacctgcgtggtggtggacgtgagccacg aggacccagaggtgaagttcaactggtacgtggacggcgtggaggtgcacaacgcca agaccaagcccagagaggagcagtacaacagcacctacagggtggtgtccgtgctga ccgtgctgcaccaggactggctgaacggcaaggaatacaagtgcaaggtctccaaca aggccctgccagcccccatcgaaaagaccatcagcaaggccaagggccagccacgg gagccccaggtgtacaccctgcccccctcccgggaggagatgaccaagaaccaggt gtccctgacctgtctggtgaagggcttctaccccagcgacatcgccgtggagtgggag agcaacggccagcccgagaacaactacaagaccacccccccagtgctggacagcga cggcagcttcttcctgtacagcaagctgaccgtggacaagtccaggtggcagcagggc aacgtgttcagctgcagcgtgatgcacgaggccctgcacaaccactacacccagaaga gcctgagcctgtcccccggcaagggaggaagcggaggatctggcggttccggaggc tcttgtgtggaacccctcggcatggaaaacggcaatatcgccaatagccagattgccgc

		cagcagcgtcagagtgacatttctgggactgcagcactgggtgcccgagctggctaga ctgaatagagccggcatggtcaacgcctggacacccagcagcaacgacgataaccctt ggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgttacacagggcgcc tctagactggccagccacgagtatctgaaggcctttaaggtggcctacagcctgaacgg ccacgagttcgacttcatccacgacgtgaacaagaagcacaaagagtttgtcggcaact ggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtggaagcccagtacgt gcggctgtaccctacaagctgtcacaccgcctgcactctgagattcgaactgctgggat gcgagctgaacggctgtgctaatcctctgggcctgaagaacaacagcatccccgataa gcagatcaccgccagctccagctataagacatggggcctgcacctgttcagctggaac ccttcttacgccagactggacaagcagggcaacttcaatgcttgggtggccggcagcta cggcaatgatcagtggctgcaagtggacctgggcagcagcaaagaagtgacaggcat catcacccagggcgccagaaatttcggcagcgtgcagtttgtggccagctacaaagtg gcctactccaacgacagcgccaactggaccgagtatcaggaccctagaaccggcagc tccaagatcttccccggcaattgggacaaccacagccacaagaagaatctgttcgaaac ccctatcctggccagatatgtgcgcattctgcccgtggcctggcacaacagaattgccct gagactggaactgctcggctgtggctctcaccaccaccatcaccat
18	FP071 EGF-Fc(u)-C1-C2	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDKTHTCPPCPAPELLGGPSVFLFPPKPKDT LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA KTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS NKALPAPIEKTISKAKGQPREPQVYTLPPCREEMTKNQ VSLWCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT QKSLSLSPGKGGSGGSGGSGGSCVEPLGMENGNIANSQ IAASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCAN PLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYARLD KQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGA RNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFP

		GNWDNHSHKKNLFETPILARYVRILPVAWHNRIALRLE LLGCGSHHHHHH
19	FP071 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgacaagacccacacctgccccccctgccca gccccagagctgctgggcggaccctccgtgttcctgttcccccccaagcccaaggaca ccctgatgatcagcaggacccccgaggtgacctgcgtggtggtggacgtgagccacg aggacccagaggtgaagttcaactggtacgtggacggcgtggaggtgcacaacgcca agaccaagcccagagaggagcagtacaacagcacctacagggtggtgtccgtgctga ccgtgctgcaccaggactggctgaacggcaaggaatacaagtgcaaggtctccaaca aggccctgccagcccccatcgaaaagaccatcagcaaggccaagggccagccacgg gagccccaggtgtacaccctgcccccctgccgggaggagatgaccaagaaccaggt gtccctgtggtgtctggtgaagggcttctaccccagcgacatcgccgtggagtgggag agcaacggccagcccgagaacaactacaagaccacccccccagtgctggacagcga cggcagcttcttcctgtacagcaagctgaccgtggacaagtccaggtggcagcagggc aacgtgttcagctgcagcgtgatgcacgaggccctgcacaaccactacacccagaaga gcctgagcctgtcccccggcaagggaggaagcggaggatctggcggttccggaggc tcttgtgtggaacccctcggcatggaaaacggcaatatcgccaatagccagattgccgc cagcagcgtcagagtgacatttctgggactgcagcactgggtgcccgagctggctaga ctgaatagagccggcatggtcaacgcctggacacccagcagcaacgacgataaccctt ggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgttacacagggcgcc tctagactggccagccacgagtatctgaaggcctttaaggtggcctacagcctgaacgg ccacgagttcgacttcatccacgacgtgaacaagaagcacaaagagtttgtcggcaact ggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtggaagcccagtacgt gcggctgtaccctacaagctgtcacaccgcctgcactctgagattcgaactgctgggat gcgagctgaacggctgtgctaatcctctgggcctgaagaacaacagcatccccgataa gcagatcaccgccagctccagctataagacatggggcctgcacctgttcagctggaac ccttcttacgccagactggacaagcagggcaacttcaatgcttgggtggccggcagcta cggcaatgatcagtggctgcaagtggacctgggcagcagcaaagaagtgacaggcat catcacccagggcgccagaaatttcggcagcgtgcagtttgtggccagctacaaagtg gcctactccaacgacagcgccaactggaccgagtatcaggaccctagaaccggcagc

		tccaagatcttccccggcaattgggacaaccacagccacaagaagaatctgttcgaaac ccctatcctggccagatatgtgcgcattctgcccgtggcctggcacaacagaattgccct gagactggaactgctcggctgtggctctcaccaccaccatcaccat
20	FP072 EGF-Fc(lõ)-C1-C2	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDKTHTCPPCPAPELLGGPSVFLFPPKPKDT LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA KTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS NKALPAPIEKTISKAKGQPREPQVCTLPPSREEMTKNQV SLSCAVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSD GSFFLVSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQ KSLSLSPGKGGSGGSGGSGGSCVEPLGMENGNIANSQI AASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCAN PLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYARLD KQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGA RNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFP GNWDNHSHKKNLFETPILARYVRILPVAWHNRIALRLE LLGCGSHHHHHH
21	FP072 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgacaagacccacacctgccccccctgccca gccccagagctgctgggcggaccctccgtgttcctgttcccccccaagcccaaggaca ccctgatgatcagcaggacccccgaggtgacctgcgtggtggtggacgtgagccacg aggacccagaggtgaagttcaactggtacgtggacggcgtggaggtgcacaacgcca agaccaagcccagagaggagcagtacaacagcacctacagggtggtgtccgtgctga ccgtgctgcaccaggactggctgaacggcaaggaatacaagtgcaaggtctccaaca aggccctgccagccccatcgaaaagaccatcagcaaggccaagggccagccacgg gagccccaggtgtgcaccctgccccctcccgggaggagatgaccaagaaccaggt gtccctgtcctgtgcggtgaagggcttctaccccagcgacatcgccgtggagtgggag

		agcaacggccagcccgagaacaactacaagaccacccccccagtgctggacagcga cggcagcttcttcctggtcagcaagctgaccgtggacaagtccaggtggcagcagggc aacgtgttcagctgcagcgtgatgcacgaggccctgcacaaccactacacccagaaga gcctgagcctgtcccccggcaagggaggaagcggaggatctggcggttccggaggc tcttgtgtggaacccctcggcatggaaaacggcaatatcgccaatagccagattgccgc cagcagcgtcagagtgacatttctgggactgcagcactgggtgcccgagctggctaga ctgaatagagccggcatggtcaacgcctggacacccagcagcaacgacgataacccctt ggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgttacacagggcgcc tctagactggccagccacgagtatctgaaggcctttaaggtggcctacagcctgaacgg ccacgagttcgacttcatccacgacgtgaacaagaagcacaaagagtttgtcggcaact ggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtggaagcccagtacgt gcggctgtaccctacaagctgtcacaccgcctgcactctgagattcgaactgctgggat gcgagctgaacggctgtgctaatcctctgggcctgaagaacaacagcatccccgataa gcagatcaccgccagctccagctataagacatggggcctgcacctgttcagctggaac ccttcttacgccagactggacaagcagggcaacttcaatgcttgggtggccggcagcta cggcaatgatcagtggctgcaagtggacctgggcagcagcaaagaagtgacaggcat catcacccagggcgccagaaatttcggcagcgtgcagtttgtggccagctacaaagtg gcctactccaacgacagcgccaactggaccgagtatcaggaccctagaaccggcagc tccaagatcttccccggcaattgggacaaccacagccacaagaagaatctgttcgaaac ccctatcctggccagatatgtgcgcattctgcccgtggcctggcacaacagaattgccct gagactggaactgctcggctgtggctctcaccaccaccatcaccat
22	FP080 EGF-C1-C2-Fc	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKCVEPLGMENGNIANSQIAASSVRVTFLGLQH WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCELNGCANPLGLKNNSIPDKQITA SSSYKTWGLHLFSWNPSYARLDKQGNFNAWVAGSYG NDQWLQVDLGSSKEVTGIITQGARNFGSVQFVASYKV AYSNDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNLF ETPILARYVRILPVAWHNRIALRLELLGCGGGGTDKTH

		TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVV DVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTY RVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISK AKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSD IAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK SRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK
23	FP080 Axit nucleic	ctggacatctgcagcaagaacccctgccacaacggcggcctgtgcgaagagatcagc caggaagtgcggggcgacgtgttccccagctacacctgtacctgcctgaagggctacg ccggcaaccactgcgagactaagtgcgtggaacccctgggcatggaaaacggcaata tcgccaacagccagatcgccgccagctccgtgcgcgtgacctttctgggactgcagca ctgggtgcccgagctggccagactgaacagagccggcatggtgaacgcctggaccc ccagcagcaacgacgacaacccttggatccaggtgaacctgctgcggcggatgtggg tgacaggcgtggtgacacagggcgccagcagactggccagccacgagtacctgaag gcctttaaggtggcctacagcctgaacggccacgagttcgacttcatccacgacgtgaa caagaaacacaaagaatttgtgggcaactggaacaagaacgccgtgcacgtgaacct gttcgagacacccgtggaagcccagtacgtgcggctgtaccccaccagctgccacac cgcctgcaccctgagattcgagctgctgggctgcgagctgaacggctgcgccaaccc cctgggcctgaagaacaacagcatccccgacaagcagatcaccgcctccagcagcta caagacctggggcctgcacctgttcagctggaaccccagctacgcccggctggacaa gcagggcaacttcaacgcctgggtggccggcagctacggcaacgaccagtggctgc aggtggacctgggcagcagcaaagaagtgaccggcatcatcacccaggggccaga aacttcggcagcgtgcagttcgtggccagctacaaagtggcctactccaacgacagcg ccaactggaccgagtaccaggacccccggaccggcagctccaagatcttccccggca actgggacaaccacagccacaagaagaatctgttcgaaacccccatcctggccagata cgtgcggatcctgcccgtggcctggcacaaccggatcgccctgagactggaactgctg ggatgtgggggaggcggtaccgacaagacccacacctgccccccctgcccagcccc agagctgctgggcggaccctccgtgttcctgttcccccccaagcccaaggacaccctg atgatcagcaggacccccgaggtgacctgcgtggtggtggacgtgagccacgaggac ccagaggtgaagttcaactggtacgtggacggcgtggaggtgcacaacgccaagacc aagcccagagaggagcagtacaacagcacctacagggtggtgtccgtgctgaccgtg ctgcaccaggactggctgaacggcaaggaatacaagtgcaaggtctccaacaaggcc

		ctgccagcccccatcgaaaagaccatcagcaaggccaagggccagccacgggagcc ccaggtgtacaccctgccccctcccgggaggagatgaccaagaaccaggtgtccct gacctgtctggtgaagggcttctaccccagcgacatcgccgtggagtgggagagcaac ggccagcccgagaacaactacaagaccacccccccagtgctggacagcgacggcag cttcttcctgtacagcaagctgaccgtggacaagtccaggtggcagcagggcaacgtgt tcagctgcagcgtgatgcacgaggccctgcacaaccactacacccagaagagcctga gcctgtcccccggcaag
24	FP090 Fc-EGF- C1-C2	DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE KTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKG FYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKL TVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK GSLEVLFQGPGSSLDICSKNPCHNGGLCEEISQEVRGDV FPSYTCTCLKGYAGNHCETKCVEPLGMENGNIANSQIA ASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSND DNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAFK VAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVNL FETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCANP LGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYARLDK QGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGAR NFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFPG NWDNHSHKKNLFETPILARYVRILPVAWHNRIALRLEL LGC
25	FP090 Axit nucleic	gacaagacccacacctgtccccctgccctgctcctgagctgctgggaggacccagc gtgttcctgttcccccccaagcccaaggacaccctgatgatcagccggacccccgaag tgacctgcgtggtggtggacgtgtcccacgaggaccctgaagtgaagttcaattggtac gtggacggcgtggaggtgcacaacgccaagaccaagccccgggaggaacagtaca acagcacctaccgggtggtgtccgtgctgaccgtgctgcaccaggactggctgaacgg caaagaatacaagtgcaaggtgtccaacaaggccctgcctgcccccatcgagaaaac catcagcaaggccaagggccagcccagagaaccccaggtgtacacactcccaccaa

		gccgggaggaaatgaccaagaaccaggtgtccctgacctgcctggtgaagggcttcta ccccagcgacattgccgtggagtgggagagcaacggccagcctgagaacaactaca agaccacccctccagtcctcgattctgatggatctttcttcctgtactccaagctgaccgtg gacaagagccggtggcagcagggaaacgtcttttcctgttccgtcatgcatgaggctct ccacaatcactacacccagaagtccctgagcctgagccccggcaagggatccctcga ggtgctgtttcagggaccaggcagcagcctggacatctgcagcaagaacccctgccac aacggcggcctgtgcgaagagatcagccaggaagtgcggggcgacgtgttccccag ctacacctgtacctgcctgaagggctacgccggcaaccactgcgagactaagtgcgtg gaacccctgggaatggaaaacggcaatatcgccaacagccagatcgccgccagctcc gtcagagtgacctttctgggactccagcactgggtgcccgagctggccagactgaata gagccggcatggtcaacgcctggacccccagcagcaacgacgacaacccctggattc aagtgaacctgctgcggcgtatgtgggtcaccggcgtcgtgacacagggcgctagca gactggccagccacgagtacctgaaggcctttaaggtggcctacagcctgaacggcca cgagttcgacttcatccacgacgtgaacaagaaacacaaagaatttgtgggcaactgga acaagaacgccgtgcacgtgaacctgttcgagacacccgtggaagcccagtacgtgc ggctgtaccctaccagctgtcacaccgcctgcaccttaagattcgagctgctgggctgc gagctgaacggctgcgctaatcctctgggcctgaagaacaacagcatccccgacaag cagatcaccgcctccagcagctacaagacctggggactgcacctgttcagctggaacc ctagctacgcccggctggacaagcagggcaacttcaatgcttgggtggccggcagcta cggcaacgaccagtggctccaggtggacctgggcagcagcaaagaagtgaccggca tcatcacccaggggccagaaacttcggcagcgtgcagttcgtggcctcctacaaagt ggcctactccaacgacagcgccaactggaccgagtaccaggaccctagaaccggca gctccaagattttccccggcaactgggataaccacagccacaagaagaatctgttcgaa acccccatcctggcccgctacgtgcgcattctaccggtcgcctggcacaaccggatcg ccctgagactggaactgctgggatgc
26	FP100 EGF-C2-C2	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGCANPLGLKNNSIPDKQITASSSYKTWGLHL FSWNPSYARLDKQGNFNAWVAGSYGNDQWLQVDLGS SKEVTGIITQGARNFGSVQFVASYKVAYSNDSANWTEY QDPRTGSSKIFPGNWDNHSHKKNLFETPILARYVRILPV AWHNRIALRLELLGCGCANPLGLKNNSIPDKQITASSSY

		KTWGLHLFSWNPSYARLDKQGNFNAWVAGSYGNDQ WLQVDLGSSKEVTGIITQGARNFGSVQFVASYKVAYSN DSANWTEYQDPRTGSSKIFPGNWDNHSHKKNLFETPIL ARYVRILPVAWHNRIALRLELLGC
27	FP100 Axit nucleic	ctggacatctgcagcaagaacccctgccacaacggcggcctgtgcgaagagatcagc caggaagtgcggggcgacgtgttccccagctacacctgtacctgcctgaagggctacg ccggcaaccactgcgagacaaagggctgcgccaacccccctgggcctgaagaacaac agcatccccgacaagcagatcaccgccagcagcagctacaagacctggggcctgca cctgttcagctggaaccccagctacgcccggctggacaagcagggcaacttcaacgc ctgggtggccggcagctacggcaacgaccagtggctgcaggtggacctgggcagca gcaaagaagtgaccggcatcatcacccagggcgccagaaacttcggcagcgtgcagt tcgtggccagctacaaggtggcctacagcaacgacagcgccaactggaccgagtacc aggacccccggaccggcagctccaagatcttccccggcaactgggacaaccacagc cacaagaagaacctgttcgagacacccatcctggccagatacgtgcggatcctgcccg tggcctggcacaaccggatcgccctgagactggaactgctgggctgcggctgtgccaa tcctctgggactgaaaaacaattccatccctgataagcagattacagcctccagctcctat aagacatgggggctgcatctgttttcttggaacccctcctacgctagactggataagcag ggaaatttcaatgcttgggtggccgggtcctatggaaatgatcagtggctgcaggtgga cctgggatcctccaaagaagtgacagggattattacacaggggctcggaactttggct ctgtgcagtttgtggcttcctacaaagtggcttactccaacgattccgccaattggacaga atatcaggatcccagaaccggctccagcaagatctttcctggaaattgggataatcactc ccacaagaaaaatctgtttgaaacccctattctggctcgctatgtgcgcattctgcctgtg gcttggcataatagaatcgctctgcggctggaactgctgggatgc
28	FP110 EGF-C1-C2-HSA	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKCVEPLGMENGNIANSQIAASSVRVTFLGLQH WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCELNGCANPLGLKNNSIPDKQITA SSSYKTWGLHLFSWNPSYARLDKQGNFNAWVAGSYG NDQWLQVDLGSSKEVTGIITQGARNFGSVQFVASYKV

		AYSNDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNLF ETPILARYVRILPVAWHNRIALRLELLGCGGSGGSGGSG GSDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQSPF EDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDK LCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNP NLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPY FYAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELR DEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRF PKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADL AKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYAR RHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKV FDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRY TKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPC AEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRP CFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKK QTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKA DDKETCFAEEGKKLVAASQAALGLHHHHHH
29	FP110 Axit nucleic	ctggacatctgcagcaagaacccctgccacaacggcggcctgtgcgaagagatcagc caggaagtgcggggcgacgtgttccccagctacacctgtacctgcctgaagggctacg ccggcaaccactgcgagactaagtgcgtggaaccccctgggcatggaaaacggcaata tcgccaacagccagatcgccgccagctccgtgcgcgtgacctttctgggactgcagca ctgggtgcccgagctggccagactgaacagagccggcatggtgaacgcctggaccc ccagcagcaacgacgacaacccttggatccaggtgaacctgctgcggcggatgtggg tgacaggcgtggtgacacagggcgccagcagactggccagccacgagtacctgaag gcctttaaggtggcctacagcctgaacggccacgagttcgacttcatccacgacgtgaa caagaaacacaaagaatttgtgggcaactggaacaagaacgccgtgcacgtgaacct gttcgagacacccgtggaagcccagtacgtgcggctgtaccccaccagctgccacac cgcctgcaccctgagattcgagctgctgggctgcgagctgaacggctgcgccaaccc cctgggcctgaagaacaacagcatccccgacaagcagatcaccgcctccagcagcta caagacctggggcctgcacctgttcagctggaaccccagctacgcccggctggacaa

		gcagggcaacttcaacgcctgggtggccggcagctacggcaacgaccagtggctgc aggtggacctgggcagcagcaaagaagtgaccggcatcatcacccaggggccaga aacttcggcagcgtgcagttcgtggccagctacaaagtggcctactccaacgacagcg ccaactggaccgagtaccaggacccccggaccggcagctccaagatcttccccggca actgggacaaccacagccacaagaagaatctgttcgaaacccccatcctggccagata cgtgcggatcctgcccgtggcctggcacaaccggatcgccctgagactggaactgctg ggatgtggaggaagcggaggatctggcggttccggaggctctgacgcccacaagag cgaggtggcccaccggttcaaggacctgggcgaggaaaacttcaaggccctggtgct gatcgccttcgcccagtacctgcagcagagccccttcgaagatcacgtaaagttagtca acgaggttacggaattcgcaaagacatgcgttgctgacgaatccgctgagaattgtgac aagagtttgcacactttattcggagataagttgtgtactgtagctactttgagagagactta cggtgaaatggctgactgctgtgcaaaacaggaaccagaacgtaacgaatgtttccttc agcataaggatgataaccctaaccttccaaggcttgttaggccagaagtcgacgtgatgt gcaccgccttccatgataatgaagagacttttcttaaaaagtacctatacgagattgcaag gcgtcatccatatttttacgccccagagctgttgtttttcgcaaagagatacaaagctgcat ttactgagtgttgccaagctgccgacaaggccgcttgtttgctaccaaagttggacgaatt gagagacgagggtaaggcatcatctgccaagcagagattaaaatgtgcatctttgcaaa aatttggagagagagcttttaaggcatgggctgttgcccgactaagccaaagattcccaa aagccgaatttgctgaagtatccaagctggtgactgatttgactaaagtacatacagaatg ttgccatggcgacctttagaatgtgctgatgacagagcagatttggctaagtatatctgc gaaaatcaagattcaatcagctctaagctgaaggaatgttgcgagaaaccactgttagaa aaatcgcattgtattgctgaagttgaaaatgatgagatgcctgctgacttgccttctcttgc cgctgattttgttgagtcgaaggatgtctgtaagaattatgctgaagctaaagacgttttcct gggtatgttcttatatgagtacgcaagacgtcacccagattactctgtggttctgctactga gattggctaaaacatacgagacaacgctggagaagtgctgtgctgccgctgaccctcat gagtgctatgcaaaggtttttgatgaattcaaaccattggttgaagagcctcaaaacttgat aaagcagaactgtgagctgtttgagcaattgggtgagtataagttccaaaatgccctgttg gtgagatatacaaaaaaggtaccccaagtttcaacgcccactttagttgaagtgtccaga aatcttggtaaagtgggtagcaaatgttgcaagcatccagaagccaagcgaatgccctg tgctgaggattatctgtccgtcgtgttgaaccaattgtgcgtattacacgaaaaaacccca gtctctgatagagtcaccaaatgttgcactgagtcactagttaatagaaggccttgttttcc gctttggaagttgatgaaacctacgtgcctaaggaatttaacgctgagacctttacctttca

		cgctgacatttgtactttgagtgaaaaagagcgtcaaatcaaaaagcaaaccgctcttgtt gaattggtgaaacacaagcctaaggctacgaaggagcagcttaaagccgtcatggacg atttcgccgcatttgttgaaaaatgctgtaaagctgatgacaaggaaacatgtttcgctga agagggaaagaaattggttgcggccagtcaggccgcacttggtttgcaccatcatcacc atcac
30	FP220 HSA-EGF-C1-C2	DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQSPFED HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC TVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNL PRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFY APELLFFAKRYKAAFTECCQAADKAACLLPKLDELRDE GKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPK AEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAK YICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPAD LPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRH PDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFD EFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTK KVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAE DYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCF SALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQT ALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADD KETCFAEEGKKLVAASQAGGSGGSGGSGGSLDICSKNP CHNGGLCEEISQEVRGDVFPSYTCTCLKGYAGNHCETK CVEPLGMENGNIANSQIAASSVRVTFLGLQHWVPELAR LNRAGMVNAWTPSSNDDNPWIQVNLLRRMWVTGVVT QGASRLASHEYLKAFKVAYSLNGHEFDFIHDVNKKHK EFVGNWNKNAVHVNLFETPVEAQYVRLYPTSCHTACT LRFELLGCELNGCANPLGLKNNSIPDKQITASSSYKTWG LHLFSWNPSYARLDKQGNFNAWVAGSYGNDQWLQVD LGSSKEVTGIITQGARNFGSVQFVASYKVAYSNDSANW TEYQDPRTGSSKIFPGNWDNHSHKKNLFETPILARYVRI LPVAWHNRIALRLELLGCGSHHHHHH

31	FP220 Axit nucleic	gacgcccacaagagcgaggtggcccaccggttcaaggacctgggcgaggaaaactt caaggccctggtgctgatcgccttcgcccagtacctgcagcagagccccttcgaagatc acgtaaagttagtcaacgaggttacggaattcgcaaagacatgcgttgctgacgaatcc gctgagaattgtgacaagagtttgcacactttattcggagataagttgtgtactgtagctac tttgagagagacttacggtgaaatggctgactgctgtgcaaaacaggaaccagaacgta acgaatgtttccttcagcataaggatgataaccctaaccttccaaggcttgttaggccaga agtcgacgtgatgtgcaccgccttccatgataatgaagagacttttcttaaaaagtacctat acgagattgcaaggcgtcatccatattttacgccccagagctgttgtttttcgcaaagag atacaaagctgcatttactgagtgttgccaagctgccgacaaggccgcttgtttgctacca aagttggacgaattgagagacgagggtaaggcatcatctgccaagcagagattaaaat gtgcatctttgcaaaaatttggagagagagcttttaaggcatgggctgttgcccgactaa gccaaagattcccaaaagccgaatttgctgaagtatccaagctggtgactgatttgacta aagtacatacagaatgttgccatggcgacctttagaatgtgctgatgacagagcagattt ggctaagtatatctgcgaaaatcaagattcaatcagctctaagctgaaggaatgttgcga gaaaccactgttagaaaaatcgcattgtattgctgaagttgaaaatgatgagatgcctgct gacttgccttctcttgccgctgattttgttgagtcgaaggatgtctgtaagaattatgctgaa gctaaagacgttttcctgggtatgttcttatatgagtacgcaagacgtcacccagattactc tgtggttctgctactgagattggctaaaacatacgagacaacgctggagaagtgctgtgc tgccgctgaccctcatgagtgctatgcaaaggtttttgatgaattcaaaccattggttgaag agcctcaaaacttgataaagcagaactgtgagctgtttgagcaattgggtgagtataagtt ccaaaatgccctgttggtgagatatacaaaaaaggtaccccaagtttcaacgcccacttt agttgaagtgtccagaaatcttggtaaagtgggtagcaaatgttgcaagcatccagaag ccaagcgaatgccctgtgctgaggattatctgtccgtcgtgttgaaccaattgtgcgtatt acacgaaaaaaccccagtctctgatagagtcaccaaatgttgcactgagtcactagttaa tagaaggccttgtttttccgctttggaagttgatgaaacctacgtgcctaaggaatttaacg ctgagacctttacctttcacgctgacatttgtactttgagtgaaaaagagcgtcaaatcaaa aagcaaaccgctcttgttgaattggtgaaacacaagcctaaggctacgaaggagcagct taaagccgtcatggacgatttcgccgcatttgttgaaaaatgctgtaaagctgatgacaa ggaaacatgtttcgctgaagagggaaagaaattggttgcggccagtcaggccggagg aagcggaggatctggcggttccggaggctctctagacatctgcagcaagaaccctgc cacaacggcggcctgtgcgaagagatcagccaggaagtgcggggcgacgtgttccc cagctacacctgtacctgcctgaagggctacgccggcaaccactgcgagactaagtgc

		gtggaacccctgggcatggaaaacggcaatatcgccaacagccagatcgccgccagc tccgtgcgcgtgacctttctgggactgcagcactgggtgcccgagctggccagactga acagagccggcatggtgaacgcctggacccccagcagcaacgacgacaacccttgg atccaggtgaacctgctgcggcggatgtgggtgacaggcgtggtgacacagggcgcc agcagactggccagccacgagtacctgaaggcctttaaggtggcctacagcctgaacg gccacgagttcgacttcatccacgacgtgaacaagaaacacaaagaatttgtgggcaa ctggaacaagaacgccgtgcacgtgaacctgttcgagacacccgtggaagcccagta cgtgcggctgtaccccaccagctgccacaccgcctgcaccctgagattcgagctgctg ggctgcgagctgaacggctgcgccaacccccctgggcctgaagaacaacagcatcccc gacaagcagatcaccgcctccagcagctacaagacctggggcctgcacctgttcagct ggaaccccagctacgcccggctggacaagcagggcaacttcaacgcctgggtggcc ggcagctacggcaacgaccagtggctgcaggtggacctgggcagcagcaaagaagt gaccggcatcatcacccaggggccagaaacttcggcagcgtgcagttcgtggccag ctacaaagtggcctactccaacgacagcgccaactggaccgagtaccaggaccccg gaccggcagctccaagatcttccccggcaactgggacaaccacagccacaagaagaa tctgttcgaaacccccatcctggccagatacgtgcggatcctgcccgtggcctggcaca accggatcgccctgagactggaactgctgggatgtggctctcaccaccaccatcaccat
32	FP250 EGF-HSA	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC

		TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGLGSHHH HHH
33	FP250 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgatgctcacaagagcgaggtggcccacagat tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaagga cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga gagatgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgcag aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct

		gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtaccct gagcgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca agaaactggtggctgcctctcaggctgctctcggacttggctctcaccaccaccatcacc at
34	FP260 EGF-HSA-C1	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGLGGSGG SGGSGGSCVEPLGMENGNIANSQIAASSVRVTFLGLQH WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCGSHHHHHH
35	FP260	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg

	Axit nucleic	gcaaccactgcgagacaaaaggatccgatgctcacaagagcgaggtggcccacagat tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaagga cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga gagatgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgcag aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtaccct gagcgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca agaaactggtggctgcctctcaggctgctctcggacttggaggaagcggaggatctgg cggttccggaggctcttgtgtggaacccctcggcatggaaaacggcaatatcgccaata

		gccagattgccgccagcagcgtcagagtgacatttctgggactgcagcactgggtgcc cgagctggctagactgaatagagccggcatggtcaacgcctggacacccagcagcaa cgacgataacccttggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgt tacacagggcgcctctagactggccagccacgagtatctgaaggcctttaaggtggcct acagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaag agtttgtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtg gaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagatt cgaactgctgggatgcggctctcaccaccaccatcaccat
36	FP270 EGF-HSA-C2	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGLGGSGG SGGSGGSCANPLGLKNNSIPDKQITASSSYKTWGLHLFS WNPSYARLDKQGNFNAWVAGSYGNDQWLQVDLGSS KEVTGIITQGARNFGSVQFVASYKVAYSNDSANWTEY QDPRTGSSKIFPGNWDNHSHKKNLFETPILARYVRILPV AWHNRIALRLELLGCGSHHHHHH

37	FP270 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgacgcccacaagagcgaggtggcccaccg gttcaaggacctgggcgaggaaaacttcaaggccctggtgctgatcgccttcgcccagt acctgcagcagagccccttcgaagatcacgtaaagttagtcaacgaggttacggaattc gcaaagacatgcgttgctgacgaatccgctgagaattgtgacaagagtttgcacacttta ttcggagataagttgtgtactgtagctactttgagagagacttacggtgaaatggctgact gctgtgcaaaacaggaaccagaacgtaacgaatgtttccttcagcataaggatgataac cctaaccttccaaggcttgttaggccagaagtcgacgtgatgtgcaccgccttccatgat aatgaagagacttttcttaaaaagtacctatacgagattgcaaggcgtcatccatatttttac gccccagagctgttgtttttcgcaaagagatacaaagctgcatttactgagtgttgccaag ctgccgacaaggccgcttgtttgctaccaaagttggacgaattgagagacgagggtaa ggcatcatctgccaagcagagattaaaatgtgcatctttgcaaaaatttggagagagagc ttttaaggcatgggctgttgcccgactaagccaaagattcccaaaagccgaatttgctga agtatccaagctggtgactgatttgactaaagtacatacagaatgttgccatggcgacctt ttagaatgtgctgatgacagagcagatttggctaagtatatctgcgaaaatcaagattcaa tcagctctaagctgaaggaatgttgcgagaaaccactgttagaaaaatcgcattgtattgc tgaagttgaaaatgatgagatgcctgctgacttgccttctcttgccgctgattttgttgagtc gaaggatgtctgtaagaattatgctgaagctaaagacgttttcctgggtatgttcttatatga gtacgcaagacgtcacccagattactctgtggttctgctactgagattggctaaaacatac gagacaacgctggagaagtgctgtgctgccgctgaccctcatgagtgctatgcaaagg tttttgatgaattcaaaccattggttgaagagcctcaaaacttgataaagcagaactgtga gctgtttgagcaattgggtgagtataagttccaaaatgccctgttggtgagatatacaaaa aaggtaccccaagtttcaacgcccactttagttgaagtgtccagaaatcttggtaaagtgg gtagcaaatgttgcaagcatccagaagccaagcgaatgccctgtgctgaggattatctgt ccgtcgtgttgaaccaattgtgcgtattacacgaaaaaaccccagtctctgatagagtca ccaaatgttgcactgagtcactagttaatagaaggccttgtttttccgctttggaagttgatg aaacctacgtgcctaaggaatttaacgctgagacctttacctttcacgctgacatttgtactt tgagtgaaaaagagcgtcaaatcaaaaagcaaaccgctcttgttgaattggtgaaacac aagcctaaggctacgaaggagcagcttaaagccgtcatggacgatttcgccgcatttgtt gaaaaatgctgtaaagctgatgacaaggaaacatgtttcgctgaagagggaaagaaatt ggttgcggccagtcaggccgcacttggtttgggaggaagcggaggatctggcggttcc

		ggaggctcttgcgccaacccccctgggcctgaagaacaacagcatccccgacaagcag atcaccgcctccagcagctacaagacctggggcctgcacctgttcagctggaacccca gctacgcccggctggacaagcagggcaacttcaacgcctgggtggccggcagctac ggcaacgaccagtggctgcaggtggacctgggcagcagcaaagaagtgaccggcat catcacccaggggccagaaacttcggcagcgtgcagttcgtggccagctacaaagt ggcctactccaacgacagcgccaactggaccgagtaccaggacccccggaccggca gctccaagatcttccccggcaactgggacaaccacagccacaagaagaatctgttcga aacccccatcctggccagatacgtgcggatcctgcccgtggcctggcacaaccggatc gccctgagactggaactgctgggatgtggctctcaccaccaccatcaccat
38	FP280 EGF(RGE)- HSA-C1-C2	LDICSKNPCHNGGLCEEISQEV**RGE**VFPSYTCTCLKGY AGNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAF AQYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDK SLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGLGGSGG SGGSGGSCVEPLGMENGNIANSQIAASSVRVTFLGLQH WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCELNGCANPLGLKNNSIPDKQITA

		SSSYKTWGLHLFSWNPSYARLDKQGNFNAWVAGSYG NDQWLQVDLGSSKEVTGIITQGARNFGSVQFVASYKV AYSNDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNLF ETPILARYVRILPVAWHNRIALRLELLGCGSHHHHHH
39	FP280 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgaggttttccccagctacacctgtacatgtctgaagggctacgcc ggcaaccactgcgagacaaaaggatccgatgctcacaagagcgaggtggcccacag attcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagt atctccagcagagcccttcgaggaccacgtgaagctggtcaacgaagtgaccgagtt cgccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacac cctgttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatg gccgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaag gacgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccg cctttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacg gcacccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttc accgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagct gagagatgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgc agaagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagattt cctaaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcaca ccgagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaa gtacatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaa gcccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgcc gatctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgc cgaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccg actattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaat gctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcct ctggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctg ggcgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcagg tgtccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgt tgcaagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgc tgaatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtg

		ctgtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgag acatacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtac cctgagcgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaa gcacaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccg cctttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagg gcaagaaactggtggctgcctctcaggctgctctcggacttggaggaagcggaggatc tggcggttccggaggctcttgtgtggaacccctcggcatggaaaacggcaatatcgcc aatagccagattgccgccagcagcgtcagagtgacatttctgggactgcagcactgggt gcccgagctggctagactgaatagagccggcatggtcaacgcctggacacccagcag caacgacgataacccttggattcaagtgaacctgctgcggcgtatgtgggtcacaggtg ttgttacacagggcgcctctagactggccagccacgagtatctgaaggcctttaaggtg gcctacagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcaca aagagtttgtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacc tgtggaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctga gattcgaactgctgggatgcgagctgaacggctgtgctaatcctctgggcctgaagaac aacagcatccccgataagcagatcaccgccagctccagctataagacatggggcctgc acctgttcagctggaacccttcttacgccagactggacaagcagggcaacttcaatgctt gggtggccggcagctacggcaatgatcagtggctgcaagtggacctgggcagcagc aaagaagtgacaggcatcatcacccagggcgccagaaatttcggcagcgtgcagtttg tggccagctacaaagtggcctactccaacgacagcgccaactggaccgagtatcagg accctagaaccggcagctccaagatcttccccggcaattgggacaaccacagccacaa gaagaatctgttcgaaacccctatcctggccagatatgtgcgcattctgcccgtggcctg gcacaacagaattgccctgagactggaactgctcggctgtggctctcaccaccaccatc accat
40	FP320 EGF-HSA D3-C1-C2- His	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSLVEEPQNLIKQNCELFEQLGEYKFQNALL VRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKR MPCAEDCLSVFLNQLCVLHEKTPVSDRVTKCCTESLVN GRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQ IKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCC KADDKETCFAEEGKKLVAASQAALGLGGSGGSGGSGG

		SCVEPLGMENGNIANSQIAASSVRVTFLGLQHWVPELA RLNRAGMVNAWTPSSNDDNPWIQVNLLRRMWVTGVV TQGASRLASHEYLKAFKVAYSLNGHEFDFIHDVNKKH KEFVGNWNKNAVHVNLFETPVEAQYVRLYPTSCHTAC TLRFELLGCELNGCANPLGLKNNSIPDKQITASSSYKTW GLHLFSWNPSYARLDKQGNFNAWVAGSYGNDQWLQV DLGSSKEVTGIITQGARNFGSVQFVASYKVAYSNDSAN WTEYQDPRTGSSKIFPGNWDNHSHKKNLFETPILARYV RILPVAWHNRIALRLELLGCGSHHHHHH
41	FP320 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccctggtggaagaaccccagaacctgatcaagc agaactgcgagctgttcgagcagctgggcgagtacaagttccagaatgccctgctcgt gcggtacaccaagaaagtgcctcaggtgtccacacctacactggtcgaggtgtccaga aacctgggcaaagtgggcagcaagtgctgcaagcaccctgaggccaaaagaatgcct tgcgccgaggattgcctgagcgtgttcctgaatcagctgtgcgtgctgcacgagaaaac ccctgtgtccgacagagtgaccaagtgctgtaccgagagcctggtcaacggcagacct tgctttagcgccctggaagtggatgagacatacgtgcccaaagagttcaacgccgaga cattcaccttccacgccgacatctgtaccctgagcgagaaagagcggcagatcaagaa gcagacagccctggtcgagctggtcaagcacaagcctaaggccaccaaagaacagct gaaggccgtgatggacgacttcgccgccttcgtggaaaagtgttgcaaggccgacgac aaagagacatgcttcgccgaagagggcaagaaactggtggctgcctctcaggctgctc tcggacttggaggaagcggaggatctggcggttccggaggctcttgtgtggaacccct cggcatggaaaacggcaatatcgccaatagccagatcgccgccagcagcgtcagagt gacatttctgggactgcagcactgggtgccagagctggctagactgaatagagccggc atggtcaacgcctggacacccagcagcaacgacgacaacccctggattcaagtgaac ctgctgcggcgtatgtgggtcacaggtgttgttacacagggcgcctctagactggccag ccacgagtatctgaaggcctttaaggtggcctacagcctgaacggccacgagttcgact tcatccacgacgtgaacaagaagcacaaagagtttgtcggcaactggaacaagaacgc cgtgcacgtgaacctgttcgagacacctgtggaagcccagtacgtgcggctgtaccta caagctgtcacaccgcctgcacactgagattcgaactgctgggatgcgagctgaacgg

		ctgtgctaatcctctgggcctgaagaacaacagcatccccgataagcagatcaccgcct ccagcagctataagacatggggcctgcacctgttcagctggaaccctagctacgccag actggacaagcagggcaactttaatgcctgggtggccggcagctacggcaatgatcaa tggctgcaagtggacctgggcagcagcaaagaagtgaccggcatcattacccagggc gcaagaaatttcggcagcgtgcagttcgtggccagctacaaagtggcctactccaacg acagcgccaactggaccgagtatcaggaccctagaaccggcagctccaagatcttccc cggcaattgggacaaccacagccacaagaagaatctgttcgaaacccctatcctggcc agatatgtgcgcattctgcccgtggcctggcacaacagaattgccctgagactggaact gctcggctgtggctctcaccaccaccatcaccat
42	FP330 EGF-HSA-C1-C2	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGLGGSGG SGGSGGSCVEPLGMENGNIANSQIAASSVRVTFLGLQH WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCELNGCANPLGLKNNSIPDKQITA

		SSSYKTWGLHLFSWNPSYARLDKQGNFNAWVAGSYG NDQWLQVDLGSSKEVTGIITQGARNFGSVQFVASYKV AYSNDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNLF ETPILARYVRILPVAWHNRIALRLELLGC
43	FP330 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgatgctcacaagagcgaggtggcccacagat tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaagga cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga gagatgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgcag aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct

		gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtaccct gagcgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca agaaactggtggctgcctctcaggctgctctcggacttggaggaagcggaggatctgg cggttccggaggctcttgtgtggaacccctcggcatggaaaacggcaatatcgccaata gccagattgccgccagcagcgtcagagtgacatttctgggactgcagcactgggtgcc cgagctggctagactgaatagagccggcatggtcaacgcctggacacccagcagcaa cgacgataacccttggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgt tacacagggcgcctctagactggccagccacgagtatctgaaggcctttaaggtggcct acagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaag agtttgtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtg gaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagatt cgaactgctgggatgcgagctgaacggctgtgctaatcctctgggcctgaagaacaac agcatccccgataagcagatcaccgccagctccagctataagacatggggcctgcacc tgttcagctggaacccttcttacgccagactggacaagcagggcaacttcaatgcttggg tggccggcagctacggcaatgatcagtggctgcaagtggacctgggcagcagcaaag aagtgacaggcatcatcacccagggcgccagaaatttcggcagcgtgcagtttgtggc cagctacaaagtggcctactccaacgacagcgccaactggaccgagtatcaggaccct agaaccggcagctccaagatcttccccggcaattgggacaaccacagccacaagaag aatctgttcgaaaccctatcctggccagatatgtgcgcattctgcccgtggcctggcac aacagaattgccctgagactggaactgctcggctgt
44	FP278 EGF-HSA-C1-C2 thẻ His	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE

		CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGLGGSGG SGGSGGSCVEPLGMENGNIANSQIAASSVRVTFLGLQH WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCELNGCANPLGLKNNSIPDKQITA SSSYKTWGLHLFSWNPSYARLDKQGNFNAWVAGSYG NDQWLQVDLGSSKEVTGIITQGARNFGSVQFVASYKV AYSNDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNLF ETPILARYVRILPVAWHNRIALRLELLGCGSHHHHHH
45	FP278 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgatgctcacaagagcgaggtggcccacagat tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaagga cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga gagatgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgcag

		aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc
		taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc
		gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt
		acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag
		cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg
		atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc
		gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga
		ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg
		ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct
		ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg
		cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt
		ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg
		caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg
		aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct
		gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac
		atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtaccct
		gagcgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc
		acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct
		ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca
		agaaactggtggctgcctctcaggctgctctcggacttggaggaagcggaggatctgg
		cggttccggaggctcttgtgtggaacccctcggcatggaaaacggcaatatcgccaata
		gccagattgccgccagcagcgtcagagtgacatttctgggactgcagcactgggtgcc
		cgagctggctagactgaatagagccggcatggtcaacgcctggacacccagcagcaa
		cgacgataacccttggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgt
		tacacagggcgcctctagactggccagccacgagtatctgaaggcctttaaggtggcct
		acagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaag
		agtttgtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtg
		gaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagatt
		cgaactgctgggatgcgagctgaacggctgtgctaatcctctgggcctgaagaacaac
		agcatccccgataagcagatcaccgccagctccagctataagacatggggcctgcacc
		tgttcagctggaacccttcttacgccagactggacaagcagggcaacttcaatgcttggg
		tggccggcagctacggcaatgatcagtggctgcaagtggacctgggcagcagcaaag

		aagtgacaggcatcatcacccagggcgccagaaatttcggcagcgtgcagtttgtggc cagctacaaagtggcctactccaacgacagcgccaactggaccgagtatcaggaccct agaaccggcagctccaagatcttccccggcaattgggacaaccacagccacaagaag aatctgttcgaaacccctatcctggccagatatgtgcgcattctgcccgtggcctggcac aacagaattgccctgagactggaactgctcggctgtggctctcaccaccaccatcacca t
46	FP068	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVAASQAALCVEPLGMENGN IANSQIAASSVRVTFLGLQHWVPELARLNRAGMVNAW TPSSNDDNPWIQVNLLRRMWVTGVVTQGASRLASHEY LKAFKVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNA VHVNLFETPVEAQYVRLYPTSCHTACTLRFELLGCELN GCANPLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSY ARLDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGII TQGARNFGSVQFVASYKVAYSNDSANWTEYQDPRTGS

		SKIFPGNWDNHSHKKNLFETPILARYVRILPVAWHNRIA LRLELLGC
47	FP068 Axit nucleic	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaggatgcccacaagagcgaggtggcccacagattcaag gatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtatctcca gcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcgccaa gacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccctgttc ggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggccgac tgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaaggacgac aaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcctttca cgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggcaccc ctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcaccgaa tgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctgagaga tgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccagaagtt cggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcctaag gccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacaccgagt gctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagtacat ctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaagcccc tgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccgatctg cctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgccgagg ccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccgactatc tgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatgctgcg ccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctctggtgg aagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagt acaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgtccaca cctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttgcaagc accctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctgaatca gctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgctgtacc gagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagacatacg tgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccctgtcc

		gagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagcacaa gcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcctttgtc gagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggcaagaa actggtggctgcttctcaggccgctctgtgtgtggaacccctcggcatggaaaacggca atatcgccaatagccagattgccgccagcagcgtcagagtgacatttctgggactgcaa cactgggtgcccgagctggctagactgaatagagccggcatggtcaacgcctggaca cccagcagcaacgacgataatccctggattcaagtgaacctgctgcggcgtatgtgggt cacaggtgttgttacacagggcgcaagcagactggccagccacgagtatctgaaggc ctttaaggtggcctacagcctgaacggccacgagttcgacttcatccacgacgtgaaca agaagcacaaagagtttgtcggcaactggaacaagaacgccgtgcacgtgaacctgtt cgagacacctgtggaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcc tgcactctgagattcgaactgctgggatgcgagctgaacggctgtgctaatcctctgggc ctgaagaacaacagcatccccgataagcagatcaccgccagctccagctataagacat ggggcctgcacctgttcagctggaacccttcttacgccagactggacaagcagggcaa cttcaatgcttgggtggccggcagctacggcaatgatcagtggctgcaagtggacctgg gcagcagcaaagaagtgacaggcatcatcacccaaggggccagaaatttcggcagc gtgcagttcgtggccagctacaaagtggcctactccaacgacagcgccaactggaccg agtatcaggaccctagaaccggcagctccaagatcttccccggcaattgggacaacca cagccacaagaagaatctgttcgaaacccctatcctggccagatatgtgcgcattctgcc cgtggcctggcacaacagaattgccctgagactggaactgctcggctgc
48	FP776	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC

		YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVACVEPLGMENGNIANSQI AASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCAN PLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYARLD KQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGA RNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFP GNWDNHSHKKNLFETPILARYVRILPVAWHNRIALRLE LLGC
49	FP776 Axit nucleic	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaggatgcccacaagagcgaggtggcccacagattcaag gatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtatctcca gcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcgccaa gacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccctgttc ggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggccgac tgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaaggacgac aaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcctttca cgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggcaccc ctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcaccgaa tgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctgagaga tgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccagaagtt cggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcctaag gccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacaccgagt gctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagtacat

		ctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaagcccc
		tgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccgatctg
		cctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgccgagg
		ccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccgactattc
		tgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatgctgcg
		ccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctctggtgg
		aagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagt
		acaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgtccaca
		cctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttgcaagc
		accctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctgaatca
		gctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgctgtacc
		gagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagacatacg
		tgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccctgtcc
		gagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagcacaa
		gcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcctttgtc
		gagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggcaagaa
		actggtggcctgtgtggaacccctcggcatggaaaacggcaatatcgccaatagccag
		attgccgccagcagcgtcagagtgacatttctgggactgcaacactgggtgcccgagct
		ggctagactgaatagagccggcatggtcaacgcctggacacccagcagcaacgacg
		ataatccctggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgttacaca
		gggcgcaagcagactggccagccacgagtatctgaaggcctttaaggtggcctacag
		cctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaagagttt
		gtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtggaa
		gcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagattcga
		actgctgggatgcgagctgaacggctgtgctaatcctctgggcctgaagaacaacagc
		atccccgataagcagatcaccgccagctccagctataagacatggggcctgcacctgtt
		cagctggaacccttcttacgccagactggacaagcagggcaacttcaatgcttgggtgg
		ccggcagctacggcaatgatcagtggctgcaagtggacctgggcagcagcaaagaa
		gtgacaggcatcatcacccaaggggccagaaatttcggcagcgtgcagttcgtggcca
		gctacaaagtggcctactccaacgacagcgccaactggaccgagtatcaggaccctag
		aaccggcagctccaagatcttccccggcaattgggacaaccacagccacaagaagaa

		tctgttcgaaacccctatcctggccagatatgtgcgcattctgcccgtggcctggcacaa cagaattgccctgagactggaactgctcggctgc
50	FP284	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETK**GS**DAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGV**GGSG** **GSGGSGGS**CVEPLGMENGNIANSQIAASSVRVTFLGLQ HWVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRR MWVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFI HDVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLY PTSCHTACTLRFELLGCELNGCANPLGLKNNSIPDKQIT ASSSYKTWGLHLFSWNPSYARLDKQGNFNAWVAGSY GNDQWLQVDLGSSKEVTGIITQGARNFGSVQFVASYK VAYSNDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNL FETPILARYVRILPVAWHNRIALRLELLGC
51	FP284 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgatgctcacaagagcgaggtggcccacagat

		tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaagga cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga gagatgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgcag aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtaccct gagcgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca agaaactggtggctgcctctcaggctgctctcggagtgggaggaagcggaggatctg gcggttccggaggctcttgtgtggaacccctcggcatggaaaacggcaatatcgccaat agccagattgccgccagcagcgtcagagtgacatttctgggactgcagcactgggtgc

		ccgagctggctagactgaatagagccggcatggtcaacgcctggacacccagcagca acgacgataacccttggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttg ttacacagggcgcctctagactggccagccacgagtatctgaaggcctttaaggtggcc tacagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaag agtttgtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtg gaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagatt cgaactgctgggatgcgagctgaacggctgtgctaatcctctgggcctgaagaacaac agcatccccgataagcagatcaccgccagctccagctataagacatggggcctgcacc tgttcagctggaacccttcttacgccagactggacaagcagggcaacttcaatgcttggg tggccggcagctacggcaatgatcagtggctgcaagtggacctgggcagcagcaaag aagtgacaggcatcatcacccagggcgccagaaatttcggcagcgtgcagtttgtggc cagctacaaagtggcctactccaacgacagcgccaactggaccgagtatcaggaccct agaaccggcagctccaagatcttccccggcaattgggacaaccacagccacaagaag aatctgttcgaaaccccctatcctggccagatatgtgcgcattctgcccgtggcctggcac aacagaattgccctgagactggaactgctcggctgt
52	FP138	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA

		FVEKCCKADDKETCFAEEGKKLVAASQAALGGSGGSG GSGGSCVEPLGMENGNIANSQIAASSVRVTFLGLQHWV PELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRMWV TGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIHDV NKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYPTS CHTACTLRFELLGCELNGCANPLGLKNNSIPDKQITASS SYKTWGLHLFSWNPSYARLDKQGNFNAWVAGSYGND QWLQVDLGSSKEVTGIITQGARNFGSVQFVASYKVAYS NDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNLFETPI LARYVRILPVAWHNRIALRLELLGC
53	FP138 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgatgctcacaagagcgaggtggcccacagat tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagcccttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaagga cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga gagatgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgcag aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg

		ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtaccct gagcgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca agaaactggtggctgcctctcaggctgctctcggaggaagcggaggatctggcggttc cggaggctcttgtgtggaacccctcggcatggaaaacggcaatatcgccaatagccag attgccgccagcagcgtcagagtgacatttctgggactgcagcactgggtgcccgagc tggctagactgaatagagccggcatggtcaacgcctggacacccagcagcaacgacg ataaccettggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgttacaca gggcgcctctagactggccagccacgagtatctgaaggcctttaaggtggcctacagc ctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaagagtttgt cggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtggaagc ccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagattcgaact gctgggatgcgagctgaacggctgtgctaatcctctgggcctgaagaacaacagcatc cccgataagcagatcaccgccagctccagctataagacatggggcctgcacctgttca gctggaacccttcttacgccagactggacaagcagggcaacttcaatgcttgggtggcc ggcagctacggcaatgatcagtggctgcaagtggacctgggcagcagcaaagaagtg acaggcatcatcacccagggcgccagaaatttcggcagcgtgcagtttgtggccagct acaaagtggcctactccaacgacagcgccaactggaccgagtatcaggaccctagaa ccggcagctccaagatcttccccggcaattgggacaaccacagccacaagaagaatct gttcgaaacccctatcctggccagatatgtgcgcattctgcccgtggcctggcacaaca gaattgccctgagactggaactgctcggctgt
54	FP811	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA

		QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGGGGSC VEPLGMENGNIANSQIAASSVRVTFLGLQHWVPELARL NRAGMVNAWTPSSNDDNPWIQVNLLRRMWVTGVVT QGASRLASHEYLKAFKVAYSLNGHEFDFIHDVNKKHK EFVGNWNKNAVHVNLFETPVEAQYVRLYPTSCHTACT LRFELLGCELNGCANPLGLKNNSIPDKQITASSSYKTWG LHLFSWNPSYARLDKQGNFNAWVAGSYGNDQWLQVD LGSSKEVTGIITQGARNFGSVQFVASYKVAYSNDSANW TEYQDPRTGSSKIFPGNWDNHSHKKNLFETPILARYVRI LPVAWHNRIALRLELLGC
55	FP811 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgatgctcacaagagcgaggtggcccacagat tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagcccttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc

		cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaagga
		cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc
		tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc
		acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac
		cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga
		gagatgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgcag
		aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc
		taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc
		gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt
		acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag
		cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg
		atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc
		gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga
		ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg
		ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct
		ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg
		cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt
		ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg
		caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg
		aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct
		gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac
		atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtaccct
		gagcgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc
		acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct
		ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca
		agaaactggtggctgcctctcaagctgctctcggaggcggaggatcttgtgtggaaccc
		ctcggcatggaaaacggcaatatcgccaatagccagattgccgccagcagcgtcaga
		gtgacatttctgggactgcagcactgggtgcccgagctggctagactgaatagagccg
		gcatggtcaacgcctggacacccagcagcaacgacgataacccttggattcaagtgaa
		cctgctgcggcgtatgtgggtcacaggtgttgttacacagggcgcctctagactggcca
		gccacgagtatctgaaggcctttaaggtggcctacagcctgaacggccacgagttcga
		cttcatccacgacgtgaacaagaagcacaaagagtttgtcggcaactggaacaagaac

		gccgtgcacgtgaacctgttcgagacacctgtggaagcccagtacgtgcggctgtacc ctacaagctgtcacaccgcctgcactctgagattcgaactgctgggatgcgagctgaac ggctgtgctaatcctctgggcctgaagaacaacagcatccccgataagcagatcaccg ccagctccagctataagacatggggcctgcacctgttcagctggaacccttcttacgcca gactggacaagcagggcaacttcaatgcttgggtggccggcagctacggcaatgatca gtggctgcaagtggacctgggcagcagcaaagaagtgacaggcatcatcacccagg gcgccagaaatttcggcagcgtgcagtttgtggccagctacaaagtggcctactccaac gacagcgccaactggaccgagtatcaggaccctagaaccggcagctccaagatcttcc ccggcaattgggacaaccacagccacaagaagaatctgttcgaaacccctatcctggc cagatatgtgcgcattctgcccgtggcctggcacaacagaattgccctgagactggaac tgctcggctgc
56	FP010	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGGGGSGG GGSCVEPLGMENGNIANSQIAASSVRVTFLGLQHWVPE LARLNRAGMVNAWTPSSNDDNPWIQVNLLRRMWVTG VVTQGASRLASHEYLKAFKVAYSLNGHEFDFIHDVNK

		KHKEFVGNWNKNAVHVNLFETPVEAQYVRLYPTSCHT ACTLRFELLGCELNGCANPLGLKNNSIPDKQITASSSYK TWGLHLFSWNPSYARLDKQGNFNAWVAGSYGNDQW LQVDLGSSKEVTGIITQGARNFGSVQFVASYKVAYSND SANWTEYQDPRTGSSKIFPGNWDNHSHKKNLFETPILA RYVRILPVAWHNRIALRLELLGC
57	FP010 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatccgatgctcacaagagcgaggtggcccacagat tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaagga cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga gagatgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgcag aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg

		caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtaccct gagcgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca agaaactggtggctgcctctcaagctgctctcggaggcggaggctccggaggcggag gatcttgtgtggaacccctcggcatggaaaacggcaatatcgccaatagccagattgcc gccagcagcgtcagagtgacatttctgggactgcagcactgggtgcccgagctggcta gactgaatagagccggcatggtcaacgcctggacacccagcagcaacgacgataacc cttggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgttacacagggcg cctctagactggccagccacgagtatctgaaggcctttaaggtggcctacagcctgaac ggccacgagttcgacttcatccacgacgtgaacaagaagcacaaagagtttgtcggca actggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtggaagcccagta cgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagattcgaactgctgg gatgcgagctgaacggctgtgctaatcctctgggcctgaagaacaacagcatccccga taagcagatcaccgccagctccagctataagacatggggcctgcacctgttcagctgga acccttcttacgccagactggacaagcagggcaacttcaatgcttgggtggccggcag ctacggcaatgatcagtggctgcaagtggacctgggcagcagcaaagaagtgacagg catcatcacccagggcgccagaaatttcggcagcgtgcagtttgtggccagctacaaa gtggcctactccaacgacagcgccaactggaccgagtatcaggaccctagaaccggc agctccaagatcttccccggcaattgggacaaccacagccacaagaagaatctgttcga aacccctatcctggccagatatgtgcgcattctgcccgtggcctggcacaacagaattg ccctgagactggaactgctcggctgc
58	FP816	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK

		LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVAASQAALGLCVEPLGMEN GNIANSQIAASSVRVTFLGLQHWVPELARLNRAGMVN AWTPSSNDDNPWIQVNLLRRMWVTGVVTQGASRLAS HEYLKAFKVAYSLNGHEFDFIHDVNKKHKEFVGNWNK NAVHVNLFETPVEAQYVRLYPTSCHTACTLRFELLGCE LNGCANPLGLKNNSIPDKQITASSSYKTWGLHLFSWNP SYARLDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVT GIITQGARNFGSVQFVASYKVAYSNDSANWTEYQDPRT GSSKIFPGNWDNHSHKKNLFETPILARYVRILPVAWHN RIALRLELLGC
59	FP816 Axit nucleic	ctggacatctgcagcaagaatccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaggatgcccacaagagcgaggtggcccacagattcaag gatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtatctcca gcagagcccttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcgccaa gacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccctgttc ggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggccgac tgctgtgccaagcaagagcccgagagaaacgagtgcttcctgcagcacaaggacgac aaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcctttca cgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggcaccc ctacttttatgccctgagctgctgttcttcgccaagcggtataaggccgccttcaccgaa

		tgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctgagaga
		tgagggcaaagccagctctgccaagcagagactgaagtgcgccagcctgcagaagtt
		cggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcctaag
		gccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacaccgagt
		gctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagtacat
		ctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaagcccc
		tgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccgatctg
		cctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgccgagg
		ccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccgactattc
		tgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatgctgcg
		ccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctctggtgg
		aagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagt
		acaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgtccaca
		cctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttgcaagc
		accctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctgaatca
		gctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgctgtacc
		gagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagacatacg
		tgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgtaccctgagc
		gagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagcacaa
		gcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcctttgtc
		gagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggcaagaa
		actggtggctgcttctcaggctgccctgggactgtgtgtggaacccctcggcatggaaa
		acggcaatatcgccaatagccagattgccgccagcagcgtcagagtgacatttctggg
		actgcagcactgggtgcccgagctggctagactgaatagagccggcatggtcaacgc
		ctggacacccagcagcaacgacgataatccctggatccaagtgaacctgctgcggcgt
		atgtgggtcacaggtgttgttacacagggcgcctctagactggccagccacgagtatct
		gaaggcctttaaggtggcctacagcctgaacggccacgagttcgacttcatccacgac
		gtgaacaagaagcacaaagagtttgtcggcaactggaacaagaacgccgtgcacgtg
		aacctgttcgagacacctgtggaagcccagtacgtgcggctgtaccctacaagctgtca
		caccgcctgcactctgagattcgaactgctgggatgcgagctgaacggctgtgctaatc
		ctctgggcctgaagaacaacagcatccccgataagcagatcaccgccagctccagcta
		taagacatggggcctgcacctgttcagctggaacccttcttacgccagactggacaagc

		agggcaacttcaatgcttgggtggccggcagctacggcaatgatcagtggctgcaagt ggacctgggcagcagcaaagaagtgacaggcatcatcacccagggcgccagaaattt cggcagcgtgcagtttgtggccagctacaaagtggcctactccaacgacagcgccaac tggaccgagtaccaggatcctagaaccggcagctccaagatcttccccggcaattggg acaaccacagccacaagaagaatctgttcgaaacccctatcctggccagatatgtgcgg attctgcccgtggcctggcacaacagaattgccctgagactggaactgctcggctgt
62	Cầu nối $(G_2S)_4$	GGSGGSGGSGGS
63	Cầu nối $(GS)_4$	GSGSGSGS
64	Cầu nối G_4S	GGGGS
65	Cầu nối $(G_4S)_2$	GGGGSGGGGS
66	GS thẻ His	GSHHHHHH
67	thẻ His	HHHHHH
68	MFG-E8 chuột	MQVSRVLAALCGMLLCASGLFAASGDFCDSSLCLNGG TCLTGQDNDIYCLCPEGFTGLVCNETERGPCSPNPCYN DAKCLVTLDTQRGDIFTEYICQCPVGYSGIHCETETNYY NLDGEYMFTTAVPNTAVPTPAPTPDLSNNLASRCSTQL GMEGGAIADSQISASSVYMGFMGLQRWGPELARLYRT GIVNAWTASNYDSKPWIQVNLLRKMRVSGVMTQGAS RAGRAEYLKTFKVAYSLDGRKFEFIQDESGGDKEFLGN LDNNSLKVNMFNPTLEAQYIKLYPVSCHRGCTLRFELL GCELHGCSEPLGLKNNTIPDSQMSASSSYKTWNLRAFG WYPHLGRLDNQGKINAWTAQSNSAKEWLQVDLGTQR QVTGIITQGARDFGHIQYVASYKVAHSDDGVQWTVYE EQGSSKVFQGNLDNNSHKKNIFEKPFMARYVRVLPVS WHNRITLRLELLGC

69	FP1776	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA NYSCECPGEFMGRNCQYKDAHKSEVAHRFKDLGEENF KALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVADE SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQ EPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNE ETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECC QAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHT ECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETT LEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCEL FEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK VGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPV SDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKA VMDDFAAFVEKCCKADDKETCFAEEGKKLVACSGPLG IEGGIISNQQITASSTHRALFGLQKWYPYYARLNKKGLI NAWTAAENDRWPWIQINLQRKMRVTGVITQGAKRIGS PEYIKSYKIAYSNDGKTWAMYKVKGTNEDMVFRGNID NNTPYANSFTPPIKAQYVRLYPQVCRRHCTLRMELLGC ELSGCSEPLGMKSGHIQDYQITASSIFRTLNMDMFTWEP RKARLDKQGKVNAWTSGHNDQSQWLQVDLLVPTKVT GIITQGAKDFGHVQFVGSYKLAYSNDGEHWTVYQDEK QRKDKVFQGNFDNDTHRKNVIDPPIYARHIRILPWSWY GRITLRSELLGC
70	FP1068	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA

		NYSCECPGEFMGRNCQYKDAHKSEVAHRFKDLGEENF KALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVADE SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQ EPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNE ETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECC QAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHT ECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETT LEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCEL FEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK VGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPV SDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKA VMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAA LCSGPLGIEGGIISNQQITASSTHRALFGLQKWYPYYAR LNKKGLINAWTAAENDRWPWIQINLQRKMRVTGVITQ GAKRIGSPEYIKSYKIAYSNDGKTWAMYKVKGTNEDM VFRGNIDNNTPYANSFTPPIKAQYVRLYPQVCRRHCTL RMELLGCELSGCSEPLGMKSGHIQDYQITASSIFRTLNM DMFTWEPRKARLDKQGKVNAWTSGHNDQSQWLQVD LLVPTKVTGIITQGAKDFGHVQFVGSYKLAYSNDGEH WTVYQDEKQRKDKVFQGNFDNDTHRKNVIDPPIYARH IRILPWSWYGRITLRSELLGC
71 =133	FP1777	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA NYSCECPGEFMGRNCQYKDAHKSEVAHRFKDLGEENF KALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVADE SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQ

		EPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNE ETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECC QAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHT ECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETT LEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCEL FEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK VGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPV SDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKA VMDDFAAFVEKCCKADDKETCFAEEGKKLVACSGPLG IEGGIISNQQITASSTHRALFGLQKWYPYYARLNKKGLI NAWTAAENDRWPWIQINLQRKMRVTGVITQGAKRIGS PEYIKSYKIAYSNDGKTWAMYKVKGTNEDMVFRGNID NNTPYANSFTPPIKAQYVRLYPQVCRRHCTLRMELLGC ELSG
72	FP1069	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA NYSCECPGEFMGRNCQYKDAHKSEVAHRFKDLGEENF KALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVADE SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQ EPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNE ETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECC QAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHT ECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETT

		LEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCEL FEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK VGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPV SDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKA VMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAA LCSGPLGIEGGIISNQQITASSTHRALFGLQKWYPYYAR LNKKGLINAWTAAENDRWPWIQINLQRKMRVTGVITQ GAKRIGSPEYIKSYKIAYSNDGKTWAMYKVKGTNEDM VFRGNIDNNTPYANSFTPPIKAQYVRLYPQVCRRHCTL RMELLGCELSG
73 =121	FP261	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVACVEPLGMENGNIANSQI AASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF

		KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNG
74 =119	FP262	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVAASQAALCVEPLGMENGN IANSQIAASSVRVTFLGLQHWVPELARLNRAGMVNAW TPSSNDDNPWIQVNLLRRMWVTGVVTQGASRLASHEY LKAFKVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNA VHVNLFETPVEAQYVRLYPTSCHTACTLRFELLGCELN G
75	MFG-E8 chiều dài đầy đủ [L76M]	MPRPRLLAALCGALLCAPSLLVALDICSKNPCHNGGLC EEISQEVRGDVFPSYTCTCLKGYAGNHCETKCVEPLGM ENGNIANSQIAASSVRVTFLGLQHWVPELARLNRAGM VNAWTPSSNDDNPWIQVNLLRRMWVTGVVTQGASRL ASHEYLKAFKVAYSLNGHEFDFIHDVNKKHKEFVGNW NKNAVHVNLFETPVEAQYVRLYPTSCHTACTLRFELLG CELNGCANPLGLKNNSIPDKQITASSSYKTWGLHLFSW

		NPSYARLDKQGNFNAWVAGSYGNDQWLQVDLGSSKE VTGIITQGARNFGSVQFVASYKVAYSNDSANWTEYQD PRTGSSKIFPGNWDNHSHKKNLFETPILARYVRILPVAW HNRIALRLELLGC
76	Miền liên kết PS MFG-E8 với [L76M]	CVEPLGMENGNIANSQIAASSVRVTFLGLQHWVPELAR LNRAGMVNAWTPSSNDDNPWIQVNLLRRMWVTGVVT QGASRLASHEYLKAFKVAYSLNGHEFDFIHDVNKKHK EFVGNWNKNAVHVNLFETPVEAQYVRLYPTSCHTACT LRFELLGCELNGCANPLGLKNNSIPDKQITASSSYKTWG LHLFSWNPSYARLDKQGNFNAWVAGSYGNDQWLQVD LGSSKEVTGIITQGARNFGSVQFVASYKVAYSNDSANW TEYQDPRTGSSKIFPGNWDNHSHKKNLFETPILARYVRI LPVAWHNRIALRLELLGC
77	Miền liên kết EGF EDIL-3 (miền giống EGF 1-2-3)	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA NYSCECPGEFMGRNCQYK
96	Miền liên kết EGF EDIL-3 (miền giống EGF 1)	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPT
97	Miền liên kết EGF EDIL-3 (miền giống EGF 2)	SAGPCTPNPCHNGGTCEISEAYRGDTFIGYVCKCPRGFN GIHCQH

98	Miền liên kết EGF EDIL-3 (miền giống EGF 3)	NINECEVEPCKNGGICTDLVANYSCECPGEFMGRNCQYK
99	Miền liên kết EGF EDIL-3 (các miền giống EGF 1 và 2)	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSVVEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFIGYVCKCPRGFNGIHCQH
100	Miền liên kết EGF EDIL-3 (miền giống EGF 2 và 3)	SAGPCTPNPCHNGGTCEISEAYRGDTFIGYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVANYSCECPGEFMGRNCQYK
101	Miền liên kết EGF EDIL-3 (miền giống EGF 1 và 3)	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSVVEVASDEEEPTNINECEVEPCKNGGICTDLVANYSCECPGEFMGRNCQYK
78	Miền liên kết PS EDIL-3	CSGPLGIEGGIISNQQITASSTHRALFGLQKWYPYYARLNKKGLINAWTAAENDRWPWIQINLQRKMRVTGVITQGAKRIGSPEYIKSYKIAYSNDGKTWAMYKVKGTNEDMVFRGNIDNNTPYANSFTPPIKAQYVRLYPQVCRRHCTLRMELLGCELSGCSEPLGMKSGHIQDYQITASSIFRTLNMDMFTWEPRKARLDKQGKVNAWTSGHNDQSQWLQVDLLVPTKVTGIITQGAKDFGHVQFVGSYKLAYSNDGEHWTVYQDEKQRKDKVFQGNFDNDTHRKNVIDPPIYARHIRILPWSWYGRITLRSELLGCTEEE

79	Miền liên kết PS EDIL-3 bị cắt cụt TEEE	CSGPLGIEGGIISNQQITASSTHRALFGLQKWYPYYARL NKKGLINAWTAAENDRWPWIQINLQRKMRVTGVITQG AKRIGSPEYIKSYKIAYSNDGKTWAMYKVKGTNEDMV FRGNIDNNTPYANSFTPPIKAQYVRLYPQVCRRHCTLR MELLGCELSGCSEPLGMKSGHIQDYQITASSIFRTLNMD MFTWEPRKARLDKQGKVNAWTSGHNDQSQWLQVDL LVPTKVTGIITQGAKDFGHVQFVGSYKLAYSNDGEHW TVYQDEKQRKDKVFQGNFDNDTHRKNVIDPPIYARHIR ILPWSWYGRITLRSELLGC
80	Miền giống EGF 1-2-3 [EDIL3]_HSA[A626-L633]bị loại bỏ_C1_C2[MFG-E8] Không có M 3163	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA NYSCECPGEFMGRNCQYKDAHKSEVAHRFKDLGEENF KALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVADE SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQ EPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNE ETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECC QAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHT ECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETT LEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCEL FEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK VGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPV SDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKA VMDDFAAFVEKCCKADDKETCFAEEGKKLVACVEPLG MENGNIANSQIAASSVRVTFLGLQHWVPELARLNRAG MVNAWTPSSNDDNPWIQVNLLRRMWVTGVVTQGASR LASHEYLKAFKVAYSLNGHEFDFIHDVNKKHKEFVGN

		WNKNAVHVNLFETPVEAQYVRLYPTSCHTACTLRFEL LGCELNGCANPLGLKNNSIPDKQITASSSYKTWGLHLFS WNPSYARLDKQGNFNAWVAGSYGNDQWLQVDLGSS KEVTGIITQGARNFGSVQFVASYKVAYSNDSANWTEY QDPRTGSSKIFPGNWDNHSHKKNLFETPILARYVRILPV AWHNRIALRLELLGC
102	Miền giống EGF 1[EDIL3]_ HSA[A626-L633]bị loại bỏ_C1_C2[MFG-E8]	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTDAHKSEVAHRFKDLGEENFKALVLIAF AQYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDK SLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVACVEPLGMENGNIA NSQIAASSVRVTFLGLQHWVPELARLNRAGMVNAWTP SSNDDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLK AFKVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVH VNLFETPVEAQYVRLYPTSCHTACTLRFELLGCELNGC ANPLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYAR LDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQ GARNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSK

		IFPGNWDNHSHKKNLFETPILARYVRILPVAWHNRIALR LELLGC
103	Miền giống EGF 2[EDIL3]_ HSA[A626-L633]bị loại bỏ_C1_C2[MFG-E8]	SAGPCTPNPCHNGGTCEISEAYRGDTFIGYVCKCPRGFN GIHCQHDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQ QSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHTLF GDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKD DNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARR HPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPKLD ELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARLS QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDR ADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEND EMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYE YARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHECY AKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALL VRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKR MPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLV NRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKER QIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKC CKADDKETCFAEEGKKLVACVEPLGMENGNIANSQIA ASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSND DNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAFK VAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVNL FETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCANP LGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYARLDK QGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGAR NFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFPG NWDNHSHKKNLFETPILARYVRILPVAWHNRIALRLEL LGC
104	Miền giống EGF 3[EDIL3]_	NINECEVEPCKNGGICTDLVANYSCECPGEFMGRNCQY KDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQSPFE DHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKL

	HSA[A626-L633]bị loại bỏ_C1_C2[MFG-E8]	CTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPN LPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYF YAPELLFFAKRYKAAFTECCQAADKAACLLPKLDELR DEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRF PKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADL AKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP ADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYAR RHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKV FDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRY TKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPC AEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRP CFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKK QTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKA DDKETCFAEEGKKLVACVEPLGMENGNIANSQIAASSV RVTFLGLQHWVPELARLNRAGMVNAWTPSSNDDNPW IQVNLLRRMWVTGVVTQGASRLASHEYLKAFKVAYSL NGHEFDFIHDVNKKHKEFVGNWNKNAVHVNLFETPVE AQYVRLYPTSCHTACTLRFELLGCELNGCANPLGLKNN SIPDKQITASSSYKTWGLHLFSWNPSYARLDKQGNFNA WVAGSYGNDQWLQVDLGSSKEVTGIITQGARNFGSVQ FVASYKVAYSNDSANWTEYQDPRTGSSKIFPGNWDNH SHKKNLFETPILARYVRILPVAWHNRIALRLELLGC
105	Miền giống EGF 1-2[EDIL3]_HSA[A626-L633]bị loại bỏ_C1_C2[MFG-E8]	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHDAHKSEVAHRFKDLGEENFK ALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVADES AENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQE PERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEE TFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQ AADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFG ERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTE

		CCHGDLLECADDRADLAKYICENQDSISSKLKECCEKP LLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYA EAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETTL EKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELF EQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK VGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPV SDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKA VMDDFAAFVEKCCKADDKETCFAEEGKKLVACVEPLG MENGNIANSQIAASSVRVTFLGLQHWVPELARLNRAG MVNAWTPSSNDDNPWIQVNLLRRMWVTGVVTQGASR LASHEYLKAFKVAYSLNGHEFDFIHDVNKKHKEFVGN WNKNAVHVNLFETPVEAQYVRLYPTSCHTACTLRFEL LGCELNGCANPLGLKNNSIPDKQITASSSYKTWGLHLFS WNPSYARLDKQGNFNAWVAGSYGNDQWLQVDLGSS KEVTGIITQGARNFGSVQFVASYKVAYSNDSANWTEY QDPRTGSSKIFPGNWDNHSHKKNLFETPILARYVRILPV AWHNRIALRLELLGC
106	Miền giống EGF 2-3[EDIL3]_HSA[A626-L633]bị loại bỏ_C1_C2[MFG-E8]	SAGPCTPNPCHNGGTCEISEAYRGDTFIGYVCKCPRGFN GIHCQHNINECEVEPCKNGGICTDLVANYSCECPGEFM GRNCQYKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL

		LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVACVEPLGMENGNIANSQI AASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCAN PLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYARLD KQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGA RNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFP GNWDNHSHKKNLFETPILARYVRILPVAWHNRIALRLE LLGC
107	Miền giống EGF 1-3[EDIL3]_HSA[A626-L633]bị loại bỏ_C1_C2[MFG-E8]	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTNINECEVEPCKNGGICTDLVANYSCECP GEFMGRNCQYKDAHKSEVAHRFKDLGEENFKALVLIA FAQYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCD KSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNE CFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKK YLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKA ACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKA WAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDL LECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHC IAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVF LGMFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAA ADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEY KFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCC KHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTK CCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADIC TLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDF

		AAFVEKCCKADDKETCFAEEGKKLVACVEPLGMENGN IANSQIAASSVRVTFLGLQHWVPELARLNRAGMVNAW TPSSNDDNPWIQVNLLRRMWVTGVVTQGASRLASHEY LKAFKVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNA VHVNLFETPVEAQYVRLYPTSCHTACTLRFELLGCELN GCANPLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSY ARLDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGII TQGARNFGSVQFVASYKVAYSNDSANWTEYQDPRTGS SKIFPGNWDNHSHKKNLFETPILARYVRILPVAWHNRIA LRLELLGC
81	Axit nucleic có Seq ID NO: 80	gacatctgcgaccccaatccttgcgagaatggcggcatttgtctgcctggactggccga tggcagcttctcttgtgaatgccccgatggcttcacagaccccaattgcagctctgtggtg gaagtggccagcgacgaggaagaacctacaagcgctggcccctgcacacccaatcc atgtcataatggcggaacctgcgagatcagcgaggcctacagaggcgataccttcatc ggctacgtgtgcaagtgccccagaggcttcaatggcatccactgccagcacaacatca acgagtgcgaggtggaaccatgcaagaacggcggcatctgtaccgacctggtggcca attactcttgcgagtgccctggcgagttcatgggcagaaactgccagtacaaggacgcc cacaagagcgaggtggcccacagattcaaggacctgggcgaagagaacttcaaggc cctggtgctgatcgccttcgctcagtatctccagcagagcccttcgaggaccacgtgaa gctggtcaacgaagtgaccgagttcgccaagacctgtggccgatgagagcgccga gaactgtgacaagagcctgcacacactgttcggcgacaagctgtgtaccgtggccaca ctgagagaaacctacggcgagatggccgactgctgtgccaagcaagagcccgagag aaacgagtgcttcctccagcacaaggatgacaaccccaacctgcctagactcgtgcgg cctgaagtggatgtgatgtgcaccgcctttcacgacaacgaggaaaccttcctgaagaa gtacctgtacgagatcgccagacggcacccctacttttatgcccctgagctgctgttcttc gccaagcggtataaggccgccttcaccgaatgttgccaggccgctgataaggctgcct gtctgctgcctaagctggacgagctgagagatgagggcaaagccagctctgccaagc agagactgaaatgcgccagcctccagaagttcggcgagagagcttttaaggcctgggc cgttgccagactgagccagagatttcctaaggccgagtttgccgaggtgtccaagctcg tgaccgatctgacaaaggtgcacaccgagtgctgtcacggcgatctgctggaatgtgcc gacgatagagccgacctggccaagtatatctgcgagaaccaggacagcatcagcagc

		aagctgaaagagtgctgcgagaagcccctgctggaaaagtctcactgtatcgccgaag
		tggaaaacgacgagatgcccgccgatctgccttctctggctgccgatttcgtggaaagc
		aaggatgtgtgcaagaactacgccgaggccaaagatgtgtttctgggcatgtttctgtat
		gagtacgcccgcagacaccccgactattctgtggttctgctgctgcggctggccaagac
		atacgagacaaccctggaaaaatgctgcgccgctgccgatcctcacgagtgttatgcca
		aggtgttcgacgagttcaagccactggtggaagaaccccagaacctgatcaagcagaa
		ctgcgagctgttcgagcagctgggcgagtacaagttccagaatgccctgctcgtgcggt
		acaccaagaaagtgcctcaggtgtccacacctacactggttgaggtgtcccggaatctg
		ggcaaagtgggcagcaagtgttgcaagcaccctgaggccaagagaatgccttgcgcc
		gaggattacctgagcgtggtgctgaatcagctgtgcgtgctgcacgagaaaaccccctgt
		gtccgacagagtgaccaagtgctgtaccgagagcctcgtgaacagaaggccttgcttta
		gcgccctggaagtggacgagacatacgtgcccaaagagttcaacgccgagacattca
		ccttccacgccgatatctgcaccctgtccgagaaagagcggcagatcaagaagcagac
		agccctggtcgagctggttaagcacaagcccaaggccaccaaagaacagctgaaggc
		cgtgatggacgacttcgccgcctttgtcgagaagtgctgcaaggccgacgacaaagag
		acatgcttcgccgaagagggcaagaaactggtggcctgtgtggaacccctcggcatgg
		aaaacggcaatatcgccaatagccagattgccgccagcagcgtcagagtgacatttctg
		ggactgcaacactgggtgcccgagctggctagactgaatagagccggcatggtcaac
		gcctggacacccagcagcaacgacgataacccctggattcaagtgaacctgctgcgg
		cgtatgtgggtcacaggtgttgttacacagggcgcaagcagactggcctctcacgagta
		cctgaaggcctttaaggtggcctacagcctgaacggccacgagttcgacttcatccacg
		acgtgaacaagaagcacaaagagtttgtcggcaactggaacaagaacgccgtgcacg
		tgaacctgttcgagacacctgtggaagcccagtacgtgcggctgtaccctacaagctgt
		cacaccgcctgcactctgagattcgaactgctgggatgcgagctgaacggctgtgctaa
		tcctctgggcctgaagaacaacagcatccccgataagcagatcaccgccagctccagc
		tataagacatggggcctgcacctgttcagctggaacccttcttacgccagactggacaa
		gcagggcaacttcaatgcttgggtggccggcagctacggcaatgatcagtggctgcaa
		gtggatctgggcagcagcaaagaagtgacaggcatcatcacccaaggggccagaaat
		ttcggcagcgtgcagttcgtggccagctacaaagtggcctactccaacgacagcgcca
		actggaccgagtatcaggaccctagaaccggcagctccaagatcttccccggcaattg
		ggacaaccacagccacaagaagaatctcttcgagactcccatcctggccagatatgtgc
		ggattctgcctgtggcctggcacaacagaatcgccctgagactggaactgctcggctgt

108	Axit nucleic có Seq ID NO: 102	gacatctgcgaccccaacccctgcgagaacggcggcatctgcctgcccggcctggcc gacggcagcttcagctgcgagtgccccgacggcttcaccgaccccaactgcagcagc gtggtggaggtggccagcgacgaggaggagcccaccgacgcccacaagagcgag gtggcccaccggttcaaggacctgggcgaggagaacttcaaggccctggtgctgatc gccttcgcccagtacctgcagcagagccccttcgaggaccacgtgaagctggtgaacg aggtgaccgagttcgccaagacctgcgtggccgacgagagcgccgagaactgcgac aagagcctgcacaccctgttcggcgacaagctgtgcaccgtggccaccctgcgggag acctacggcgagatggccgactgctgcgccaagcaggagcccgagcggaacgagtg cttcctgcagcacaaggacgacaaccccaacctgccccggctggtgcggcccgaggt ggacgtgatgtgcaccgccttccacgacaacgaggagaccttcctgaagaagtacctg tacgagatcgcccggcggcacccctacttctacgcccccgagctgctgttcttcgccaa gcggtacaaggccgccttcaccgagtgctgccaggccgccgacaaggccgcctgcct gctgcccaagctggacgagctgcgggacgagggcaaggccagcagcgccaagcag cggctgaagtgcgccagcctgcagaagttcggcgagcgggccttcaaggcctgggcc gtggcccggctgagccagcggttccccaaggccgagttcgccgaggtgagcaagctg gtgaccgacctgaccaaggtgcacaccgagtgctgccacggcgacctgctggagtgc gccgacgaccgggccgacctggccaagtacatctgcgagaaccaggacagcatcag cagcaagctgaaggagtgctgcgagaagcccctgctggagaagagccactgcatcgc cgaggtggagaacgacgagatgcccgccgacctgcccagcctggccgccgacttcgt ggagagcaaggacgtgtgcaagaactacgccgaggccaaggacgtgttcctgggcat gttcctgtacgagtacgcccggcggcaccccgactacagcgtggtgctgctgctgcgg ctggccaagacctacgagaccaccctggagaagtgctgcgccgccgccgaccccca cgagtgctacgccaaggtgttcgacgagttcaagcccctggtggaggagccccagaa cctgatcaagcagaactgcgagctgttcgagcagctgggcgagtacaagttccagaac gccctgctggtgcggtacaccaagaaggtgccccaggtgagcacccccaccctggtg gaggtgagccggaacctgggcaaggtgggcagcaagtgctgcaagcaccccgagg ccaagcggatgccctgcgccgaggactacctgagcgtggtgctgaaccagctgtgcgt gctgcacgagaagacccccgtgagcgaccgggtgaccaagtgctgcaccgagagcc tggtgaaccggcggccctgcttcagcgccctggaggtggacgagacctacgtgccca aggagttcaacgccgagaccttcaccttccacgccgacatctgcaccctgagcgagaa ggagcggcagatcaagaagcagaccgccctggtggagctggtgaagcacaagccca aggccaccaaggagcagctgaaggccgtgatggacgacttcgccgccttcgtggaga

		agtgctgcaaggccgacgacaaggagacctgcttcgccgaggagggcaagaagctg gtggcctgcgtggagcccctgggcatggagaacggcaacatcgccaacagccagatc gccgccagcagcgtgcgggtgaccttcctgggcctgcagcactgggtgcccgagctg gcccggctgaaccgggccggcatggtgaacgcctggacccccagcagcaacgacg acaacccctggatccaggtgaacctgctgcggcggatgtgggtgaccggcgtggtga cccagggcgccagccggctggccagccacgagtacctgaaggccttcaaggtggcct acagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaagg agttcgtgggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacccccgt ggaggcccagtacgtgcggctgtaccccaccagctgccacaccgcctgcaccctgcg gttcgagctgctgggctgcgagctgaacggctgcgccaacccccctgggcctgaagaa caacagcatccccgacaagcagatcaccgccagcagcagctacaagacctggggcct gcacctgttcagctggaaccccagctacgcccggctggacaagcagggcaacttcaa cgcctgggtggccggcagctacggcaacgaccagtggctgcaggtggacctgggca gcagcaaggaggtgaccggcatcatcacccagggcgcccggaacttcggcagcgtg cagttcgtggccagctacaaggtggcctacagcaacgacagcgccaactggaccgag taccaggacccccggaccggcagcagcaagatcttccccggcaactgggacaacca cagccacaagaagaacctgttcgagacccccatcctggcccggtacgtgcggatcctg cccgtggcctggcacaaccggatcgccctgcggctggagctgctgggctgc
109	Axit nucleic có Seq ID NO: 103	agcgccggcccctgcacccccaaccccctgccacaacggcggcacctgcgagatcag cgaggcctaccggggcgacaccttcatcggctacgtgtgcaagtgcccccggggcttc aacggcatccactgccagcacgacgcccacaagagcgaggtggcccaccggttcaa ggacctgggcgaggagaacttcaaggccctggtgctgatcgccttcgcccagtacctg cagcagagcccccttcgaggaccacgtgaagctggtgaacgaggtgaccgagttcgcc aagacctgcgtggccgacgagagcgccgagaactgcgacaagagcctgcacaccct gttcggcgacaagctgtgcaccgtggccaccctgcgggagacctacggcgagatggc cgactgctgcgccaagcaggagcccgagcggaacgagtgcttcctgcagcacaagg acgacaaccccaacctgccccggctggtgcggcccgaggtggacgtgatgtgcaccg ccttccacgacaacgaggagaccttcctgaagaagtacctgtacgagatcgcccggcg gcacccctacttctacgcccccgagctgctgttcttcgccaagcggtacaaggccgcctt caccgagtgctgccaggccgccgacaaggccgcctgcctgctgcccaagctggacg agctgcgggacgagggcaaggccagcagcgccaagcagcggctgaagtgcgccag

		cctgcagaagttcggcgagcgggccttcaaggcctgggccgtggcccggctgagcca
		gcggttccccaaggccgagttcgccgaggtgagcaagctggtgaccgacctgaccaa
		ggtgcacaccgagtgctgccacggcgacctgctggagtgcgccgacgaccgggccg
		acctggccaagtacatctgcgagaaccaggacagcatcagcagcaagctgaaggagt
		gctgcgagaagcccctgctggagaagagccactgcatcgccgaggtggagaacgac
		gagatgcccgccgacctgcccagcctggccgccgacttcgtggagagcaaggacgt
		gtgcaagaactacgccgaggccaaggacgtgttcctgggcatgttcctgtacgagtac
		gcccggcggcaccccgactacagcgtggtgctgctgctgcggctggccaagacctac
		gagaccaccctggagaagtgctgcgccgccgccgacccccacgagtgctacgccaa
		ggtgttcgacgagttcaagcccctggtggaggagccccagaacctgatcaagcagaa
		ctgcgagctgttcgagcagctgggcgagtacaagttccagaacgccctgctggtgcgg
		tacaccaagaaggtgccccaggtgagcacccccaccctggtggaggtgagccggaa
		cctgggcaaggtgggcagcaagtgctgcaagcaccccgaggccaagcggatgccct
		gcgccgaggactacctgagcgtggtgctgaaccagctgtgcgtgctgcacgagaaga
		cccccgtgagcgaccgggtgaccaagtgctgcaccgagagcctggtgaaccggcgg
		ccctgcttcagcgccctggaggtggacgagacctacgtgcccaaggagttcaacgcc
		gagaccttcaccttccacgccgacatctgcaccctgagcgagaaggagcggcagatc
		aagaagcagaccgccctggtggagctggtgaagcacaagcccaaggccaccaagga
		gcagctgaaggccgtgatggacgacttcgccgccttcgtggagaagtgctgcaaggc
		cgacgacaaggagacctgcttcgccgaggagggcaagaagctggtggcctgcgtgg
		agcccctgggcatggagaacggcaacatcgccaacagccagatcgccgccagcagc
		gtgcgggtgaccttcctgggcctgcagcactgggtgcccgagctggcccggctgaac
		cgggccggcatggtgaacgcctggacccccagcagcaacgacgacaaccccctggat
		ccaggtgaacctgctgcggcggatgtgggtgaccggcgtggtgacccagggcgcca
		gccggctggccagccacgagtacctgaaggccttcaaggtggcctacagcctgaacg
		gccacgagttcgacttcatccacgacgtgaacaagaagcacaaggagttcgtgggcaa
		ctggaacaagaacgccgtgcacgtgaacctgttcgagacccccgtggaggcccagta
		cgtgcggctgtaccccaccagctgccacaccgcctgcaccctgcggttcgagctgctg
		ggctgcgagctgaacggctgcgccaaccccctgggcctgaagaacaacagcatcccc
		gacaagcagatcaccgccagcagcagctacaagacctggggcctgcacctgttcagc
		tggaaccccagctacgcccggctggacaagcagggcaacttcaacgcctgggtggcc
		ggcagctacggcaacgaccagtggctgcaggtggacctgggcagcagcaaggaggt

		gaccggcatcatcacccagggcgcccggaacttcggcagcgtgcagttcgtggccag ctacaaggtggcctacagcaacgacagcgccaactggaccgagtaccaggaccccc ggaccggcagcagcaagatcttccccggcaactgggacaaccacagccacaagaag aacctgttcgagacccccatcctggcccggtacgtgcggatcctgcccgtggcctggc acaaccggatcgccctgcggctggagctgctgggctgc
110	Axit nucleic có Seq ID NO: 104	aacatcaacgagtgcgaggtggagccctgcaagaacggcggcatctgcaccgacctg gtggccaactacagctgcgagtgccccggcgagttcatgggccggaactgccagtac aaggacgcccacaagagcgaggtggcccaccggttcaaggacctgggcgaggaga acttcaaggccctggtgctgatcgccttcgcccagtacctgcagcagagccccttcgag gaccacgtgaagctggtgaacgaggtgaccgagttcgccaagacctgcgtggccgac gagagcgccgagaactgcgacaagagcctgcacaccctgttcggcgacaagctgtgc accgtggccaccctgcgggagacctacggcgagatggccgactgctgcgccaagca ggagcccgagcggaacgagtgcttcctgcagcacaaggacgacaaccccaacctgc cccggctggtgcggcccgaggtggacgtgatgtgcaccgccttccacgacaacgagg agaccttcctgaagaagtacctgtacgagatcgcccggcggcacccctacttctacgcc cccgagctgctgttcttcgccaagcggtacaaggccgccttcaccgagtgctgccagg ccgccgacaaggccgcctgcctgctgcccaagctggacgagctgcgggacgagggc aaggccagcagcgccaagcagcggctgaagtgcgccagcctgcagaagttcggcga gcgggccttcaaggcctgggccgtggcccggctgagccagcggttccccaaggccg agttcgccgaggtgagcaagctggtgaccgacctgaccaaggtgcacaccgagtgct gccacggcgacctgctggagtgcgccgacgaccgggccgacctggccaagtacatct gcgagaaccaggacagcatcagcagcaagctgaaggagtgctgcgagaagcccctg ctggagaagagccactgcatcgccgaggtggagaacgacgagatgcccgccgacct gcccagcctggccgccgacttcgtggagagcaaggacgtgtgcaagaactacgccga ggccaaggacgtgttcctgggcatgttcctgtacgagtacgcccggcggcaccccgac tacagcgtggtgctgctgctgcggctggccaagacctacgagaccaccctggagaagt gctgcgccgccgccgaccccccacgagtgctacgccaaggtgttcgacgagttcaagc ccctggtggaggagccccagaacctgatcaagcagaactgcgagctgttcgagcagc tgggcgagtacaagttccagaacgccctgctggtgcggtacaccaagaaggtgcccc aggtgagcaccccccaccctggtggaggtgagccggaacctgggcaaggtgggcagc aagtgctgcaagcaccccgaggccaagcggatgccctgcgccgaggactacctgag

		cgtggtgctgaaccagctgtgcgtgctgcacgagaagacccccgtgagcgaccgggt gaccaagtgctgcaccgagagcctggtgaaccggcggccctgcttcagcgccctgga ggtggacgagacctacgtgcccaaggagttcaacgccgagaccttcaccttccacgcc gacatctgcaccctgagcgagaaggagcggcagatcaagaagcagaccgccctggt ggagctggtgaagcacaagcccaaggccaccaaggagcagctgaaggccgtgatgg acgacttcgccgccttcgtggagaagtgctgcaaggccgacgacaaggagacctgctt cgccgaggagggcaagaagctggtggcctgcgtggagcccctgggcatggagaac ggcaacatcgccaacagccagatcgccgccagcagcgtgcgggtgaccttcctgggc ctgcagcactgggtgcccgagctggcccggctgaaccgggccggcatggtgaacgc ctggacccccagcagcaacgacgacaacccctggatccaggtgaacctgctgcggcg gatgtgggtgaccggcgtggtgacccagggcgccagccggctggccagccacgagt acctgaaggccttcaaggtggcctacagcctgaacggccacgagttcgacttcatccac gacgtgaacaagaagcacaaggagttcgtgggcaactggaacaagaacgccgtgca cgtgaacctgttcgagacccccgtggaggcccagtacgtgcggctgtaccccaccagc tgccacaccgcctgcaccctgcggttcgagctgctgggctgcgagctgaacggctgcg ccaacccccctgggcctgaagaacaacagcatccccgacaagcagatcaccgccagc agcagctacaagacctggggcctgcacctgttcagctggaaccccagctacgcccgg ctggacaagcagggcaacttcaacgcctgggtggccggcagctacggcaacgacca gtggctgcaggtggacctgggcagcagcaaggaggtgaccggcatcatcacccagg gcgcccggaacttcggcagcgtgcagttcgtggccagctacaaggtggcctacagca acgacagcgccaactggaccgagtaccaggacccccggaccggcagcagcaagat cttccccggcaactgggacaaccacagccacaagaagaacctgttcgagacccccatc ctggcccggtacgtgcggatcctgcccgtggcctggcacaaccggatcgccctgcgg ctggagctgctgggctgc
111	Axit nucleic có Seq ID NO: 105	gacatctgcgaccccaacccctgcgagaacggcggcatctgcctgcccggcctggcc gacggcagcttcagctgcgagtgccccgacggcttcaccgaccccaactgcagcagc gtggtggaggtggccagcgacgaggaggagcccaccagcgccggcccctgcaccc ccaacccctgccacaacggcggcacctgcgagatcagcgaggcctaccggggcgac accttcatcggctacgtgtgcaagtgccccccggggcttcaacggcatccactgccagca cgacgcccacaagagcgaggtggcccaccggttcaaggacctgggcgaggagaact tcaaggcctggtgctgatcgccttcgcccagtacctgcagcagagcccccttcgagga

		ccacgtgaagctggtgaacgaggtgaccgagttcgccaagacctgcgtggccgacga gagcgccgagaactgcgacaagagcctgcacaccctgttcggcgacaagctgtgcac cgtggccaccctgcgggagacctacggcgagatggccgactgctgcgccaagcagg agcccgagcggaacgagtgcttcctgcagcacaaggacgacaaccccaacctgccc cggctggtgcggcccgaggtggacgtgatgtgcaccgccttccacgacaacgaggag accttcctgaagaagtacctgtacgagatcgcccggcggcacccctacttctacgcccc cgagctgctgttcttcgccaagcggtacaaggccgccttcaccgagtgctgccaggcc gccgacaaggccgcctgcctgctgcccaagctggacgagctgcgggacgagggcaa ggccagcagcgccaagcagcggctgaagtgcgccagcctgcagaagttcggcgagc gggccttcaaggcctgggccgtggcccggctgagccagcggttccccaaggccgagt tcgccgaggtgagcaagctggtgaccgacctgaccaaggtgcacaccgagtgctgcc acggcgacctgctggagtgcgccgacgaccgggccgacctggccaagtacatctgc gagaaccaggacagcatcagcagcaagctgaaggagtgctgcgagaagcccctgct ggagaagagccactgcatcgccgaggtggagaacgacgagatgcccgccgacctgc ccagcctggccgccgacttcgtggagagcaaggacgtgtgcaagaactacgccgag gccaaggacgtgttcctgggcatgttcctgtacgagtacgcccggcggcaccccgact acagcgtggtgctgctgctgcggctggccaagacctacgagaccaccctggagaagt gctgcgccgccgccgacccccacgagtgctacgccaaggtgttcgacgagttcaagc ccctggtggaggagccccagaacctgatcaagcagaactgcgagctgttcgagcagc tgggcgagtacaagttccagaacgccctgctggtgcggtacaccaagaaggtgcccc aggtgagcacccccaccctggtggaggtgagccggaacctgggcaaggtgggcagc aagtgctgcaagcaccccgaggccaagcggatgccctgcgccgaggactacctgag cgtggtgctgaaccagctgtgcgtgctgcacgagaagacccccgtgagcgaccgggt gaccaagtgctgcaccgagagcctggtgaaccggcggccctgcttcagcgccctgga ggtggacgagacctacgtgcccaaggagttcaacgccgagaccttcaccttccacgcc gacatctgcaccctgagcgagaaggagcggcagatcaagaagcagaccgccctggt ggagctggtgaagcacaagcccaaggccaccaaggagcagctgaaggccgtgatgg acgacttcgccgccttcgtggagaagtgctgcaaggccgacgacaaggagacctgctt cgccgaggagggcaagaagctggtggcctgcgtggagccccctgggcatggagaac ggcaacatcgccaacagccagatcgccgccagcagcgtgcgggtgaccttcctgggc ctgcagcactgggtgcccgagctggcccggctgaaccgggccggcatggtgaacgc ctggaccccagcagcaacgacgacaacccctggatccaggtgaacctgctgcggcg

		gatgtgggtgaccggcgtggtgacccagggcgccagccggctggccagccacgagt acctgaaggccttcaaggtggcctacagcctgaacggccacgagttcgacttcatccac gacgtgaacaagaagcacaaggagttcgtgggcaactggaacaagaacgccgtgca cgtgaacctgttcgagacccccgtggaggcccagtacgtgcggctgtaccccaccagc tgccacaccgcctgcaccctgcggttcgagctgctgggctgcgagctgaacggctgcg ccaacccccctgggcctgaagaacaacagcatccccgacaagcagatcaccgccagc agcagctacaagacctggggcctgcacctgttcagctggaaccccagctacgcccgg ctggacaagcagggcaacttcaacgcctgggtggccggcagctacggcaacgacca gtggctgcaggtggacctgggcagcagcaaggaggtgaccggcatcatcacccagg gcgcccggaacttcggcagcgtgcagttcgtggccagctacaaggtggcctacagca acgacagcgccaactggaccgagtaccaggacccccggaccggcagcagcaagat cttccccggcaactgggacaaccacagccacaagaagaacctgttcgagacccccatc ctggcccggtacgtgcggatcctgcccgtggcctggcacaaccggatcgccctgcgg ctggagctgctgggctgc
112	Axit nucleic có Seq ID NO: 106	agcgccggcccctgcacccccaaccccctgccacaacggcggcacctgcgagatcag cgaggcctaccggggcgacaccttcatcggctacgtgtgcaagtgcccccggggcttc aacggcatccactgccagcacaacatcaacgagtgcgaggtggagccctgcaagaac ggcggcatctgcaccgacctggtggccaactacagctgcgagtgccccggcgagttc atgggccggaactgccagtacaaggacgcccacaagagcgaggtggcccaccggtt caaggacctgggcgaggagaacttcaaggccctggtgctgatcgccttcgcccagtac ctgcagcagagccccttcgaggaccacgtgaagctggtgaacgaggtgaccgagttc gccaagacctgcgtggccgacgagagcgccgagaactgcgacaagagcctgcacac cctgttcggcgacaagctgtgcaccgtggccaccctgcgggagacctacggcgagat ggccgactgctgcgccaagcaggagcccgagcggaacgagtgcttcctgcagcaca aggacgacaaccccaacctgccccggctggtgcggcccgaggtggacgtgatgtgc accgccttccacgacaacgaggagaccttcctgaagaagtacctgtacgagatcgccc ggcggcaccctacttctacgccccgagctgctgttcttcgccaagcggtacaaggcc gccttcaccgagtgctgccaggccgccgacaaggccgcctgcctgctgcccaagctg gacgagctgcgggacgagggcaaggccagcagcgccaagcagcggctgaagtgc gccagcctgcagaagttcggcgagcgggccttcaaggcctgggccgtggcccggct gagccagcggttccccaaggccgagttcgccgaggtgagcaagctggtgaccgacct

		gaccaaggtgcacaccgagtgctgccacggcgacctgctggagtgcgccgacgacc gggccgacctggccaagtacatctgcgagaaccaggacagcatcagcagcaagctg aaggagtgctgcgagaagcccctgctggagaagagccactgcatcgccgaggtgga gaacgacgagatgcccgccgacctgcccagcctggccgccgacttcgtggagagca aggacgtgtgcaagaactacgccgaggccaaggacgtgttcctgggcatgttcctgta cgagtacgcccggcggcaccccgactacagcgtggtgctgctgcggctggccaa gacctacgagaccaccctggagaagtgctgcgccgccgccgacccccacgagtgcta cgccaaggtgttcgacgagttcaagcccctggtggaggagccccagaacctgatcaa gcagaactgcgagctgttcgagcagctgggcgagtacaagttccagaacgccctgctg gtgcggtacaccaagaaggtgccccaggtgagcacccccaccctggtggaggtgag ccggaacctgggcaaggtgggcagcaagtgctgcaagcaccccgaggccaagcgg atgccctgcgccgaggactacctgagcgtggtgctgaaccagctgtgcgtgctgcacg agaagacccccgtgagcgaccgggtgaccaagtgctgcaccgagagcctggtgaac cggcggccctgcttcagcgccctggaggtggacgagacctacgtgcccaaggagttc aacgccgagaccttcaccttccacgccgacatctgcaccctgagcgagaaggagcgg cagatcaagaagcagaccgccctggtggagctggtgaagcacaagcccaaggccac caaggagcagctgaaggccgtgatggacgacttcgccgccttcgtggagaagtgctg caaggccgacgacaaggagacctgcttcgccgaggagggcaagaagctggtggcct gcgtggagcccctgggcatggagaacggcaacatcgccaacagccagatcgccgcc agcagcgtgcgggtgaccttcctgggcctgcagcactgggtgcccgagctggcccgg ctgaaccgggccggcatggtgaacgcctggaccccagcagcaacgacgacaaccc ctggatccaggtgaacctgctgcggcggatgtgggtgaccggcgtggtgacccaggg cgccagccggctggccagccacgagtacctgaaggccttcaaggtggcctacagcct gaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaggagttcgtg ggcaactggaacaagaacgccgtgcacgtgaacctgttcgagaccccgtggaggcc cagtacgtgcggctgtaccccaccagctgccacaccgcctgcaccctgcggttcgagc tgctgggctgcgagctgaacggctgcgccaacccccctgggcctgaagaacaacagca tccccgacaagcagatcaccgccagcagcagctacaagacctggggcctgcacctgtt cagctggaaccccagctacgcccggctggacaagcagggcaacttcaacgcctgggt ggccggcagctacggcaacgaccagtggctgcaggtggacctgggcagcagcaag gaggtgaccggcatcatcacccagggcgcccggaacttcggcagcgtgcagttcgtg gccagctacaaggtggcctacagcaacgacagcgccaactggaccgagtaccagga

		ccccggaccggcagcagcaagatcttccccggcaactgggacaaccacagccaca agaagaacctgttcgagacccccatcctggcccggtacgtgcggatcctgcccgtggc ctggcacaaccggatcgccctgcggctggagctgctgggctgc
113	Axit nucleic có Seq ID NO: 107	gacatctgcgaccccaacccctgcgagaacggcggcatctgcctgcccggcctggcc gacggcagcttcagctgcgagtgccccgacggcttcaccgaccccaactgcagcagc gtggtggaggtggccagcgacgaggaggagcccaccaacatcaacgagtgcgaggt ggagccctgcaagaacggcggcatctgcaccgacctggtggccaactacagctgcga gtgccccggcgagttcatgggccggaactgccagtacaaggacgcccacaagagcg aggtggcccaccggttcaaggacctgggcgaggagaacttcaaggccctggtgctga tcgccttcgcccagtacctgcagcagagccccttcgaggaccacgtgaagctggtgaa cgaggtgaccgagttcgccaagacctgcgtggccgacgagagcgccgagaactgcg acaagagcctgcacaccctgttcggcgacaagctgtgcaccgtggccaccctgcggg agacctacggcgagatggccgactgctgcgccaagcaggagcccgagcggaacga gtgcttcctgcagcacaaggacgacaaccccaacctgccccggctggtgcggcccga ggtggacgtgatgtgcaccgccttccacgacaacgaggagaccttcctgaagaagtac ctgtacgagatcgcccggcggcacccctacttctacgcccccgagctgctgttcttcgc caagcggtacaaggccgccttcaccgagtgctgccaggccgccgacaaggccgcct gcctgctgcccaagctggacgagctgcgggacgagggcaaggccagcagcgccaa gcagcggctgaagtgcgccagcctgcagaagttcggcgagcgggccttcaaggcctg ggccgtggcccggctgagccagcggttccccaaggccgagttcgccgaggtgagca agctggtgaccgacctgaccaaggtgcacaccgagtgctgccacggcgacctgctgg agtgcgccgacgaccgggccgacctggccaagtacatctgcgagaaccaggacagc atcagcagcaagctgaaggagtgctgcgagaagcccctgctggagaagagccactgc atcgccgaggtggagaacgacgagatgcccgccgacctgcccagcctggccgccga cttcgtggagagcaaggacgtgtgcaagaactacgccgaggccaaggacgtgttcctg ggcatgttcctgtacgagtacgcccggcggcaccccgactacagcgtggtgctgctgc tgcggctggccaagacctacgagaccaccctggagaagtgctgcgccgccgccgac ccccacgagtgctacgccaaggtgttcgacgagttcaagcccctggtggaggagccc cagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagtacaagttcc agaacgccctgctggtgcggtacaccaagaaggtgccccaggtgagcacccccaccc tggtggaggtgagccggaacctgggcaaggtgggcagcaagtgctgcaagcacccc

		gaggccaagcggatgccctgcgccgaggactacctgagcgtggtgctgaaccagctg tgcgtgctgcacgagaagacccccgtgagcgaccgggtgaccaagtgctgcaccga gagcctggtgaaccggcggccctgcttcagcgccctggaggtggacgagacctacgt gcccaaggagttcaacgccgagaccttcaccttccacgccgacatctgcaccctgagc gagaaggagcggcagatcaagaagcagaccgccctggtggagctggtgaagcacaa gcccaaggccaccaaggagcagctgaaggccgtgatggacgacttcgccgccttcgt ggagaagtgctgcaaggccgacgacaaggagacctgcttcgccgaggagggcaag aagctggtggcctgcgtggagcccctgggcatggagaacggcaacatcgccaacag ccagatcgccgccagcagcgtgcgggtgaccttcctgggcctgcagcactgggtgcc cgagctggcccggctgaaccgggccggcatggtgaacgcctggacccccagcagca acgacgacaacccctggatccaggtgaacctgctgcggcggatgtgggtgaccggcg tggtgacccagggcgccagccggctggccagccacgagtacctgaaggccttcaag gtggcctacagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaagc acaaggagttcgtgggcaactggaacaagaacgccgtgcacgtgaacctgttcgaga cccccgtggaggcccagtacgtgcggctgtaccccaccagctgccacaccgcctgca ccctgcggttcgagctgctgggctgcgagctgaacggctgcgccaacccccctgggcct gaagaacaacagcatccccgacaagcagatcaccgccagcagcagctacaagacct ggggcctgcacctgttcagctggaaccccagctacgcccggctggacaagcagggca acttcaacgcctgggtggccggcagctacggcaacgaccagtggctgcaggtggacc tgggcagcagcaaggaggtgaccggcatcatcacccagggcgcccggaacttcggc agcgtgcagttcgtggccagctacaaggtggcctacagcaacgacagcgccaactgg accgagtaccaggacccccggaccggcagcagcaagatcttccccggcaactggga caaccacagccacaagaagaacctgttcgagacccccatcctggcccggtacgtgcg gatcctgcccgtggcctggcacaaccggatcgccctgcggctggagctgctgggctg c
82	EGF[MFG-E8]_HSA[A626-L633]bị loại bỏ_C1_C2[EDIL3]	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK

		LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVACSGPLGIEGGIISNQQITA SSTHRALFGLQKWYPYYARLNKKGLINAWTAAENDR WPWIQINLQRKMRVTGVITQGAKRIGSPEYIKSYKIAYS NDGKTWAMYKVKGTNEDMVFRGNIDNNTPYANSFTP PIKAQYVRLYPQVCRRHCTLRMELLGCELSGCSEPLGM KSGHIQDYQITASSIFRTLNMDMFTWEPRKARLDKQGK VNAWTSGHNDQSQWLQVDLLVPTKVTGIITQGAKDFG HVQFVGSYKLAYSNDGEHWTVYQDEKQRKDKVFQGN FDNDTHRKNVIDPPIYARHIRILPWSWYGRITLRSELLG C
83	Axit nucleic có Seq ID NO: 82	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaggatgcccacaagagcgaggtggcccacagattcaag gatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtatctcca gcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcgccaa gacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccctgttc ggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggccgac tgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaaggacgac aaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcctttca cgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggcaccc ctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcaccgaa

		tgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctgagaga
		tgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccagaagtt
		cggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcctaag
		gccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacaccgagt
		gctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagtacat
		ctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaagcccc
		tgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccgatctg
		cctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgccgagg
		ccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccgactattc
		tgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatgctgcg
		ccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctctggtgg
		aagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagt
		acaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgtccaca
		cctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttgcaagc
		accctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctgaatca
		gctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgctgtacc
		gagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagacatacg
		tgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccctgtcc
		gagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagcacaa
		gcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcctttgtc
		gagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggcaagaa
		actggtggcctgttctggccctctgggcatcgaaggcggcatcatcagcaatcagcaga
		tcaccgccagcagcacccacagagcactgtttggcctgcaaaagtggtatccctactac
		gcccggctgaacaagaagggcctgattaacgcctggacagccgccgagaatgacag
		atggccctggattcagatcaacctccagcggaagatgagagtgaccggcgttatcaca
		cagggcgcaaagagaatcggctcccctgagtacatcaagagctacaagatcgcctaca
		gcaacgacggcaagacctgggccatgtacaaagtgaagggcaccaacgaggacatg
		gtgttccggggcaacatcgacaacaacacccccttacgccaacagcttcacccctcctat
		caaggcccagtacgtgcggctgtaccctcaagtgtgcagaaggcactgtaccctgaga
		atggaactgctgggctgcgaactgtctggctgttctgagccactgggaatgaagtccgg
		ccacatccaggactaccagattaccgcctccagcatcttcagaaccctgaacatggatat
		gttcacctgggagccccggaaggccagactggataagcagggaaaagtgaatgcctg

		gaccagcggccacaacgaccagtctcaatggctgcaagtggacctgctggtgcctacc aaagtgaccggaatcatcacccaaggcgctaaggatttcggccacgtgcagttcgtgg gctcctacaagctggcctactccaatgatggcgagcactggaccgtgtaccaggacga gaagcagcggaaggataaggtgttccagggaaacttcgataacgatacccaccggaa gaacgtgatcgaccctccaatctacgccagacacatcagaatcctgccttggtcttggta cggcagaatcaccctgagatccgagctgctgggatgc
115	EGF-C1-His6	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKCVEPLGMENGNIANSQIAASSVRVTFLGLQH WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCELNGHHHHHH
116	Axit nucleic của 115	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaatgtgtggaacccctcggcatggaaaacggcaatatcg ccaatagccagatcgccgccagcagcgtcagagtgacatttctgggactgcaacactg ggtgccagagctggccagactgaatagagccggcatggttaacgcctggacacccag cagcaacgacgacaaccccctggattcaagtgaacctgctgcggcgtatgtgggtcaca ggtgttgttacacagggcgcaagcagactggccagccacgagtatctgaaggcctttaa ggtggcctacagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaag cacaaagagttcgtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgaga cacctgtggaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcac actgagattcgagctgctgggctgtgaactgaatggccaccaccaccatcaccac
117 =147	EGF-HSA-C1	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW

		AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGLGGSGG SGGSGGSCVEPLGMENGNIANSQIAASSVRVTFLGLQH WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCELNG
118	Axit nucleic của 117	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatctgatgcccacaagagcgaggtggcccacagat tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaagga cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga gagatgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccag aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt

		acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccc tgtccgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca agaaactggtggctgcctctcaggctgctctcggacttggtggaagcggaggaagtgg tggatctggcggatcttgtgtggaacccctcggcatggaaaacggcaatatcgccaata gccagattgccgccagcagcgtcagagtgacatttctgggactgcaacactgggtgcc cgagctggctagactgaatagagccggcatggtcaacgcctggacacccagcagcaa cgacgataatccctggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgt tacacagggcgcaagcagactggccagccacgagtatctgaaggcctttaaggtggc ctacagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaa gagtttgtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgt ggaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagat tcgaactgctgggatgcgagctgaacggc
119 = 74	EGF-HSA-C1	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA

		RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVAASQAALCVEPLGMENGN IANSQIAASSVRVTFLGLQHWVPELARLNRAGMVNAW TPSSNDDNPWIQVNLLRRMWVTGVVTQGASRLASHEY LKAFKVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNA VHVNLFETPVEAQYVRLYPTSCHTACTLRFELLGCELN G
120	Axit nucleic của 119	ctggacatctgtagcaagaaccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaggatgcccacaagagcgaggtggcccacagattcaag gatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtatctcca gcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcgccaa gacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccctgttc ggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggccgac tgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaaggacgac aaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcctttca cgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggcaccc ctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcaccgaa tgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctgagaga tgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccagaagtt cggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcctaag

		gccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacaccgagt gctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagtacat ctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaagcccc tgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccgatctg cctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgccgagg ccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccgactattc tgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatgctgcg ccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctctggtgg aagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagt acaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgtccaca cctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttgcaagc accctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctgaatca gctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgctgtacc gagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagacatacg tgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccctgtcc gagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagcacaa gcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcctttgtc gagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggcaagaa actggtggctgcttctcaggccgctctgtgtgtggaacccctcggcatggaaaacggca atatcgccaatagccagattgccgccagcagcgtcagagtgacatttctgggactgcaa cactgggtgcccgagctggctagactgaatagagccggcatggtcaacgcctggaca cccagcagcaacgacgataatccctggattcaagtgaacctgctgcggcgtatgtgggt cacaggtgttgttacacagggcgcaagcagactggccagccacgagtatctgaaggc ctttaaggtggcctacagcctgaacggccacgagttcgacttcatccacgacgtgaaca agaagcacaaagagtttgtcggcaactggaacaagaacgccgtgcacgtgaacctgtt cgagacacctgtggaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcc tgcactctgagattcgaactgctgggatgcgagctgaacggc
121 = 73	FP135	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH

		KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVACVEPLGMENGNIANSQI AASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNG
122	Axit nucleic của 121	ctggacatctgtagcaagaaccettgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaggatgcccacaagagcgaggtggcccacagattcaag gatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtatctcca gcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcgccaa gacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccctgttc ggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggccgac tgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaaggacgac aaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcctttca cgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggcaccc ctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcaccgaa tgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctgagaga tgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccagaagtt cggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcctaag

		gccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacaccgagt gctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagtacat ctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaagcccc tgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccgatctg cctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgccgagg ccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccgactattc tgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatgctgcg ccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctctggtgg aagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagt acaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgtccaca cctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttgcaagc accctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctgaatca gctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgctgtacc gagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagacatacg tgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccctgtcc gagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagcacaa gcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcctttgtc gagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggcaagaa actggtggcctgtgtggaacccctcggcatggaaaacggcaatatcgccaatagccag attgccgccagcagcgtcagagtgacatttctgggactgcaacactgggtgcccgagct ggctagactgaatagagccggcatggtcaacgcctggacacccagcagcaacgacg ataatccctggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgttacaca gggcgcaagcagactggccagccacgagtatctgaaggcctttaaggtggcctacag cctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaagagttt gtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtggaa gcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagattcga actgctgggatgcgagctgaacggc
123	hIgG1_FC_DAPA_Hole	DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT CVVVAVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALAAPIE KTISKAKGQPREPQVCTLPPSRDELTKNQVSLSCAVKGF

		YPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLVSKLT VDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK
124	Axit nucleic của 123	gataagacccacacctgtcctccatgtcctgctccagaactgctcggcggaccctccgtt ttcctgtttccacctaagcctaaggacaccctgatgatcagcagaacccctgaagtgacc tgtgtggtggtggccgtgtctcacgaagatcccgaagtgaagttcaattggtacgtggac ggcgtggaagtgcacaacgccaagaccaagcctagagaggaacagtacaacagcac ctacagagtggtgtccgtgctgaccgtgctgcaccaggattggctgaacggcaaagag tacaagtgcaaggtgtccaacaaggccctggccgctcctatcgagaaaaccatctctaa ggccaagggccagcctcgggaacctcaagtctgtacactgcctcctagccgggacga gctgaccaaaaatcaggtgtccctgagctgcgccgtgaagggcttttacccttccgatat cgccgtggaatgggagagcaatggccagcctgagaacaactacaagaccacacctcc tgtgctggacagcgacggctcattctttctggtgtccaagctgacagtggacaagagca gatggcagcagggcaacgtgttcagctgttctgtgatgcacgaggccctgcacaacca ctacacccagaagtctctgtctctgagccccggcaaa
125	EGF_hIgG1_FC_DAPA_Knob_C1	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDKTHTCPPCPAPELLGGPSVFLFPPKPKDT LMISRTPEVTCVVVAVSHEDPEVKFNWYVDGVEVHNA KTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS NKALAAPIEKTISKAKGQPREPQVYTLPPCREEMTKNQ VSLWCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT QKSLSLSPGKGGSGGSGGSGGSCVEPLGMENGNIANSQ IAASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCAN PLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYARLD KQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGA RNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFP GNWDNHSHKKNLFETPILARYVRILPVAWHNRIALRLE LLGC

126	Axit nucleic của 125	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaagggcagcgataagacccacacctgtcctccatgtcctgc tccagaactgctcggcggaccctccgttttcctgtttccacctaagcctaaggacaccct gatgatcagcagaacccctgaagtgacctgtgtggtggtggccgtgtctcacgaagatc ccgaagtgaagttcaattggtacgtggacggcgtggaagtgcacaacgccaagacca agcctagagaggaacagtacaacagcacctacagagtggtgtccgtgctgaccgtgct gcaccaggattggctgaacggcaaagagtacaagtgcaaggtgtccaacaaggccct ggccgctcctatcgagaaaaccatctctaaggccaagggccagcctcgggaacctca ggtttacaccctgcctccatgccgggaagagatgaccaagaatcaggtgtccctgtggt gcctggtcaagggcttctacccttccgatatcgccgtggaatgggagagcaatggcca gcctgagaacaactacaagaccacacctcctgtgctggacagcgacggctcattcttcc tgtacagcaagctgacagtggacaagagcagatggcagcagggcaacgtgttcagct gttctgtgatgcacgaggccctgcacaaccactacacccagaagtctctgtctctgagcc ctggcaaaggcggaagcggtggaagcggaggatctggcggatcttgtgtggaacccc tcggcatggaaaacggcaatatcgccaatagccagatcgccgccagcagcgtcagag tgacatttctgggactgcaacactgggtgccagagctggccagactgaatagagccgg catggttaacgcctggacacccagcagcaacgacgacaacccctggattcaagtgaac ctgctgcggcgtatgtgggtcacaggtgttgttacacagggcgcaagcagactggcca gccacgagtatctgaaggcctttaaggtggcctacagcctgaacggccacgagttcga cttcatccacgacgtgaacaagaagcacaaagagttcgtcggcaactggaacaagaac gccgtgcacgtgaacctgttcgagacacctgtggaagcccagtacgtgcggctgtacc ctacaagctgtcacaccgcctgcacactgagattcgagctgctgggctgcgagctgaat ggctgtgctaatcctctgggcctgaagaacaatagcatccccgacaagcagatcaccg cctccagcagctataagacatggggcctgcacctgtttagctggaaccctagctacgcc agactggacaagcagggaaacttcaatgcctgggtggccggcagctacggcaatgat caatggctgcaagtggacctgggcagcagcaaagaagtgaccggcatcattacccag ggcgctagaaatttcggcagcgtgcagttcgtggccagctacaaagtggcctactccaa cgacagcgccaactggaccgagtatcaggaccctagaaccggcagctccaagatcttc cccggcaattgggacaaccacagccacaagaagaatctgttcgaaaccctatcctgg

		ccagatatgtgcgcattctgcccgtggcctggcacaacagaattgccctgagactggaa ctgctgggatgc
127	hIgG1_FC_ DAPA_Kno b	DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT CVVVAVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALAAPIE KTISKAKGQPREPQVYTLPPCREEMTKNQVSLWCLVK GFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS KLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP GK
128	Axit nucleic của 127	gataagacccacacctgtcctccatgtcctgctccagaactgctcggcggaccctccgtt ttcctgtttccacctaagcctaaggacaccctgatgatcagcagaacccctgaagtgacc tgtgtggtggtggccgtgtctcacgaagatcccgaagtgaagttcaattggtacgtggac ggcgtggaagtgcacaacgccaagaccaagcctagagaggaacagtacaacagcac ctacagagtggtgtccgtgctgaccgtgctgcaccaggattggctgaacggcaaagag tacaagtgcaaggtgtccaacaaggccctggccgctcctatcgagaaaaccatctctaa ggccaagggccagcctcgggaacctcaggtttacaccctgcctccatgccgggaaga gatgaccaagaatcaggtgtccctgtggtgcctggtcaagggcttctacccttccgatat cgccgtggaatgggagagcaatggccagcctgagaacaactacaagaccacacctcc tgtgctggacagcgacggctcattcttcctgtacagcaagctgacagtggacaagagca gatggcagcagggcaacgtgttcagctgttctgtgatgcacgaggccctgcacaacca ctacacccagaagtctctgtctctgagccccggcaaa
129	EGF_hIgG1 _FC_DAPA _Hole_C1	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDKTHTCPPCPAPELLGGPSVFLFPPKPKDT LMISRTPEVTCVVVAVSHEDPEVKFNWYVDGVEVHNA KTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS NKALAAPIEKTISKAKGQPREPQVCTLPPSRDELTKNQV SLSCAVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSD GSFFLVSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQ KSLSLSPGKGGSGGSGGSGGSCVEPLGMENGNIANSQI AASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN

		DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCAN PLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYARLD KQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIITQGA RNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFP GNWDNHSHKKNLFETPILARYVRILPVAWHNRIALRLE LLGC
130	Axit nucleic của 129	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaagggcagcgataagacccacacctgtcctccatgtcctgc tccagaactgctcggcggaccctccgttttcctgtttccacctaagcctaaggacaccct gatgatcagcagaacccctgaagtgacctgtgtggtggtggccgtgtctcacgaagatc ccgaagtgaagttcaattggtacgtggacggcgtggaagtgcacaacgccaagacca agcctagagaggaacagtacaacagcacctacagagtggtgtccgtgctgaccgtgct gcaccaggattggctgaacggcaaagagtacaagtgcaaggtgtccaacaaggccct ggccgctcctatcgagaaaaccatctctaaggccaagggccagcctcgggaacctcaa gtctgtacactgcctcctagccgggacgagctgaccaaaaatcaggtgtccctgagctg cgccgtgaagggcttttacccttccgatatcgccgtggaatgggagagcaatggccag cctgagaacaactacaagaccacacctcctgtgctggacagcgacggctcattctttctg gtgtccaagctgacagtggacaagagcagatggcagcagggcaacgtgttcagctgtt ctgtgatgcacgaggccctgcacaaccactacacccagaagtctctgtctctgagccct ggcaaaggcggaagcggtggaagcggaggatctggcggatcttgtgtggaaccccctc ggcatggaaaacggcaatatcgccaatagccagatcgccgccagcagcgtcagagtg acatttctgggactgcaacactgggtgccagagctggccagactgaatagagccggca tggttaacgcctggacacccagcagcaacgacgacaacccctggattcaagtgaacct gctgcggcgtatgtgggtcacaggtgttgttacacagggcgcaagcagactggccagc cacgagtatctgaaggcctttaaggtggcctacagcctgaacggccacgagttcgactt catccacgacgtgaacaagaagcacaaagagttcgtcggcaactggaacaagaacgc cgtgcacgtgaacctgttcgagacacctgtggaagcccagtacgtgcggctgtaccta caagctgtcacaccgcctgcacactgagattcgagctgctgggctgcgagctgaatgg

		ctgtgctaatcctctgggcctgaagaacaatagcatccccgacaagcagatcaccgcct ccagcagctataagacatggggcctgcacctgtttagctggaaccctagctacgccaga ctggacaagcagggaaacttcaatgcctgggtggccggcagctacggcaatgatcaat ggctgcaagtggacctgggcagcagcaaagaagtgaccggcatcattacccagggc gctagaaatttcggcagcgtgcagttcgtggccagctacaaagtggcctactccaacga cagcgccaactggaccgagtatcaggaccctagaaccggcagctccaagatcttcccc ggcaattgggacaaccacagccacaagaagaatctgttcgaaacccctatcctggcca gatatgtgcgcattctgcccgtggcctggcacaacagaattgccctgagactggaactg ctgggatgc
131	EGF(RGE) _HSA[A62 6-L633]bị loại bỏ_C1	LDICSKNPCHNGGLCEEISQEVRGEVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVACVEPLGMENGNIANSQI AASSVRVTFLGLQHWVPELARLNRAGMVNAWTPSSN DDNPWIQVNLLRRMWVTGVVTQGASRLASHEYLKAF KVAYSLNGHEFDFIHDVNKKHKEFVGNWNKNAVHVN LFETPVEAQYVRLYPTSCHTACTLRFELLGCELNG

132	Axit nucleic của 131	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatcagtc aagaagtgcggggcgaagtctttcccagctacacctgtacctgtctgaagggctatgcc ggcaaccactgcgagacaaaggatgcccacaagagcgaggtggcccacagattcaa ggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtatctcc agcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcgccaa gacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccctgttc ggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggccgac tgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaaggacgac aaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcctttca cgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggcaccc ctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcaccgaa tgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctgagaga tgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccagaagtt cggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcctaag gccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacaccgagt gctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagtacat ctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaagcccc tgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccgatctg cctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgccgagg ccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccgactattc tgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatgctgcg ccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctctggtgg aagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagt acaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgtccaca cctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttgcaagc accctgaggccaagagaatgccttgcgccgaggattatctgagcgtggtgctgaatca gctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgctgtacc gagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagacatacg tgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccctgtcc gagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagcacaa gcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcctttgtc gagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggcaagaa

		actggtggcctgtgtggaacccctcggcatggaaaacggcaatatcgccaatagccag attgccgccagcagcgtcagagtgacatttctgggactgcaacactgggtgcccgagct ggctagactgaatagagccggcatggtcaacgcctggacacccagcagcaacgacg ataatccctggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgttacaca gggcgcaagcagactggccagccacgagtatctgaaggcctttaaggtggcctacag cctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaagagttt gtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgtggaa gcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagattcga actgctgggatgcgagctgaacggc
133 =71	EGF[EDIL3]_HSA[A62 6-L633]bị loại bỏ_C1[EDI L3]	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA NYSCECPGEFMGRNCQYKDAHKSEVAHRFKDLGEENF KALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVADE SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQ EPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNE ETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECC QAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHT ECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETT LEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCEL FEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK VGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPV SDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKA VMDDFAAFVEKCCKADDKETCFAEEGKKLVACSGPLG IEGGIISNQQITASSTHRALFGLQKWYPYYARLNKKGLI NAWTAAENDRWPWIQINLQRKMRVTGVITQGAKRIGS PEYIKSYKIAYSNDGKTWAMYKVKGTNEDMVFRGNID

		NNTPYANSFTPPIKAQYVRLYPQVCRRHCTLRMELLGC ELSG
134	Axit nucleic của 133	gacatctgcgaccccaatccttgcgagaatggcggcatttgtctgcctggactggccga tggcagcttctcttgtgaatgccccgatggcttcacagaccccaattgcagctctgtggtg gaagtggccagcgacgaggaagaacctacaagcgctggcccctgcacacccaatcc atgtcataatggcggaacctgcgagatcagcgaggcctacagaggcgataccttcatc ggctacgtgtgcaagtgccccagaggcttcaatggcatccactgccagcacaacatca acgagtgcgaggtggaaccatgcaagaacggcggcatctgtaccgacctggtggcca attactcttgcgagtgccctggcgagttcatgggcagaaactgccagtacaaggacgcc cacaagagcgaggtggcccacagattcaaggacctgggcgaagagaacttcaaggc cctggtgctgatcgccttcgctcagtatctccagcagagccctttcgaggaccacgtgaa gctggtcaacgaagtgaccgagttcgccaagacctgtgtggccgatgagagcgccga gaactgtgacaagagcctgcacacactgttcggcgacaagctgtgtaccgtggccaca ctgagagaaacctacggcgagatggccgactgctgtgccaagcaagagcccgagag aaacgagtgcttcctccagcacaaggatgacaaccccaacctgcctagactcgtgcgg cctgaagtggatgtgatgtgcaccgcctttcacgacaacgaggaaaccttcctgaagaa gtacctgtacgagatcgccagacggcacccctacttttatgcccctgagctgctgttcttc gccaagcggtataaggccgccttcaccgaatgttgccaggccgctgataaggctgcct gtctgctgcctaagctggacgagctgagagatgagggcaaagccagctctgccaagc agagactgaaatgcgccagcctccagaagttcggcgagagagcttttaaggcctgggc cgttgccagactgagccagagatttcctaaggccgagtttgccgaggtgtccaagctcg tgaccgatctgacaaaggtgcacaccgagtgctgtccggcgatctgctggaatgtgcc gacgatagagccgacctggccaagtatatctgcgagaaccaggacagcatcagcagc aagctgaaagagtgctgcgagaagcccctgctggaaaagtctcactgtatcgccgaag tggaaaacgacgagatgcccgccgatctgccttctctggctgccgatttcgtggaaagc aaggatgtgtgcaagaactacgccgaggccaaagatgtgtttctgggcatgtttctgtat gagtacgcccgcagacaccccgactattctgtggttctgctgctgcggctggccaagac atacgagacaaccctggaaaaatgctgcgccgctgccgatcctcacgagtgttatgcca aggtgttcgacgagttcaagccactggtggaagaaccccagaacctgatcaagcagaa ctgcgagctgttcgagcagctgggcgagtacaagttccagaatgccctgctcgtgcggt acaccaagaaagtgcctcaggtgtccacacctacactggttgaggtgtcccggaatctg

		ggcaaagtgggcagcaagtgttgcaagcaccctgaggccaagagaatgccttgcgcc gaggattacctgagcgtggtgctgaatcagctgtgcgtgctgcacgagaaaaccccctgt gtccgacagagtgaccaagtgctgtaccgagagcctcgtgaacagaaggccttgcttta gcgccctggaagtggacgagacatacgtgcccaaagagttcaacgccgagacattca ccttccacgccgatatctgcaccctgtccgagaaagagcggcagatcaagaagcagac agccctggtcgagctggttaagcacaagcccaaggccaccaaagaacagctgaaggc cgtgatggacgacttcgccgcctttgtcgagaagtgctgcaaggccgacgacaaagag acatgcttcgccgaagagggcaagaaactggtggcctgttctggccctctgggcatcg aaggcggcatcatcagcaatcagcagatcaccgccagcagcacccacagagcactgt ttggcctgcaaaagtggtatccctactacgcccggctgaacaagaagggcctgattaac gcctggacagccgccgagaatgacagatggccctggattcagatcaacctccagcgg aagatgagagtgaccggcgttatcacacagggcgcaaagagaatcggctcccctgag tacatcaagagctacaagatcgcctacagcaacgacggcaagacctgggccatgtaca aagtgaagggcaccaacgaggacatggtgttccggggcaacatcgacaacaacaccc cttacgccaacagcttcacccctcctatcaaggcccagtacgtgcggctgtaccctcaa gtgtgcagaaggcactgtaccctgagaatggaactgctgggctgcgaactgtctggc
135	EGF[EDIL3]_HSA[A626-L633]bỏ loại bỏ_C2[EDIL3]	DICDPNPCENGGICLPGLADGSFSCECPDGFTDPNCSSV VEVASDEEEPTSAGPCTPNPCHNGGTCEISEAYRGDTFI GYVCKCPRGFNGIHCQHNINECEVEPCKNGGICTDLVA NYSCECPGEFMGRNCQYKDAHKSEVAHRFKDLGEENF KALVLIAFAQYLQQSPFEDHVKLVNEVTEFAKTCVADE SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQ EPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNE ETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECC QAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKF GERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHT ECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNY AEAKDVFLGMFLYEYARRHPDYSVVLLLRLAKTYETT LEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCEL FEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK

		VGSKCCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPV SDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKA VMDDFAAFVEKCCKADDKETCFAEEGKKLVACSEPLG MKSGHIQDYQITASSIFRTLNMDMFTWEPRKARLDKQG KVNAWTSGHNDQSQWLQVDLLVPTKVTGIITQGAKDF GHVQFVGSYKLAYSNDGEHWTVYQDEKQRKDKVFQG NFDNDTHRKNVIDPPIYARHIRILPWSWYGRITLRSELL GC
136	Axit nucleic của 135	gacatctgcgaccccaatccttgcgagaatggcggcatttgtctgcctggactggccga tggcagcttctcttgtgaatgccccgatggcttcacagaccccaattgcagctctgtggtg gaagtggccagcgacgaggaagaacctacaagcgctggcccctgcacacccaatcc atgtcataatggcggaacctgcgagatcagcgaggcctacagaggcgataccttcatc ggctacgtgtgcaagtgccccagaggcttcaatggcatccactgccagcacaacatca acgagtgcgaggtggaaccatgcaagaacggcggcatctgtaccgacctggtggcca attactcttgcgagtgccctggcgagttcatgggcagaaactgccagtacaaggacgcc cacaagagcgaggtggcccacagattcaaggacctgggcgaagagaacttcaaggc cctggtgctgatcgccttcgctcagtatctccagcagagccctttcgaggaccacgtgaa gctggtcaacgaagtgaccgagttcgccaagacctgtgtggccgatgagagcgccga gaactgtgacaagagcctgcacacactgttcggcgacaagctgtgtaccgtggccaca ctgagagaaacctacggcgagatggccgactgctgtgccaagcaagagcccgagag aaacgagtgcttcctccagcacaaggatgacaaccccaacctgcctagactcgtgcgg cctgaagtggatgtgatgtgcaccgcctttcacgacaacgaggaaaccttcctgaagaa gtacctgtacgagatcgccagacggcacccctacttttatgcccctgagctgctgttcttc gccaagcggtataaggccgccttcaccgaatgttgccaggccgctgataaggctgcct gtctgctgcctaagctggacgagctgagagatgagggcaaagccagctctgccaagc agagactgaaatgcgccagcctccagaagttcggcgagagagcttttaaggcctgggc cgttgccagactgagccagagatttcctaaggccgagtttgccgaggtgtccaagctcg tgaccgatctgacaaaggtgcacaccgagtgctgtcacggcgatctgctggaatgtgcc gacgatagagccgacctggccaagtatatctgcgagaaccaggacagcatcagcagc aagctgaaagagtgctgcgagaagcccctgctggaaaagtctcactgtatcgccgaag

		tggaaaacgacgagatgcccgccgatctgccttctctggctgccgatttcgtggaaagc aaggatgtgtgcaagaactacgccgaggccaaagatgtgtttctgggcatgtttctgtat gagtacgcccgcagacaccccgactattctgtggttctgctgctgcggctggccaagac atacgagacaaccctggaaaaatgctgcgccgctgccgatcctcacgagtgttatgcca aggtgttcgacgagttcaagccactggtggaagaaccccagaacctgatcaagcagaa ctgcgagctgttcgagcagctgggcgagtacaagttccagaatgccctgctcgtgcggt acaccaagaaagtgcctcaggtgtccacacctacactggttgaggtgtcccggaatctg ggcaaagtgggcagcaagtgttgcaagcaccctgaggccaagagaatgccttgcgcc gaggattacctgagcgtggtgctgaatcagctgtgcgtgctgcacgagaaaacccctgt gtccgacagagtgaccaagtgctgtaccgagagcctcgtgaacagaaggccttgcttta gcgccctggaagtggacgagacatacgtgcccaaagagttcaacgccgagacattca ccttccacgccgatatctgcaccctgtccgagaaagagcggcagatcaagaagcagac agccctggtcgagctggttaagcacaagcccaaggccaccaaagaacagctgaaggc cgtgatggacgacttcgccgcctttgtcgagaagtgctgcaaggccgacgacaaagag acatgcttcgccgaagagggcaagaaactggtggcctgttctgagccactgggcatga agtctggccacatccaggattaccagatcaccgccagcagcatcttcagaaccctgaac atggatatgttcacctgggagccccggaaggccagactggataagcagggaaaagtg aacgcctggaccagcggccacaatgaccagtctcagtggctgcaagtggacctgctg gtgcctaccaaagtgaccggcatcatcacacagggcgcaaaggatttcggccacgtgc agtttgtgggcagctacaagctggcctacagcaacgatggcgagcactggacagtgta ccaggacgagaagcagcggaaggataaggtgttccagggcaacttcgacaacgaca cccaccggaagaacgtgatcgaccctcctatctacgcccggcacatcagaatcctgcct tggtcttggtacggccggatcaccctgagaagcgagctgcttggatgt
137	EGF_HSA[A626- L633]bị loại bỏ_C2	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD

		RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVACANPLGLKNNSIPDKQIT ASSSYKTWGLHLFSWNPSYARLDKQGNFNAWVAGSY GNDQWLQVDLGSSKEVTGIITQGARNFGSVQFVASYK VAYSNDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNL FETPILARYVRILPVAWHNRIALRLELLGC
138	Axit nucleic của 137	ctggacatctgtagcaagaaccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaggatgcccacaagagcgaggtggcccacagattcaag gatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtatctcca gcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcgccaa gacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccctgttc ggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggccgac tgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaaggacgac aaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcctttca cgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggcaccc ctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcaccgaa tgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctgagaga tgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccagaagtt cggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcctaag gccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacaccgagt gctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagtacat ctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaagcccc tgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccgatctg

		cctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgccgagg ccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccgactattc tgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatgctgcg ccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctctggtgg aagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagt acaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgtccaca cctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttgcaagc accctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctgaatca gctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgctgtacc gagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagacatacg tgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccctgtcc gagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagcacaa gcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcctttgtc gagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggcaagaa actggtggcctgtgctaaccctctgggcctgaagaacaacagcatccccgataagcag atcaccgccagcagcagctataagacatggggcctgcacctgttcagctggaacccttc ttacgccagactggacaagcagggcaacttcaatgcttgggtggccggcagctacggc aatgatcagtggctgcaagtggacctgggcagcagcaaagaagtgacaggcatcatc acccagggcgcaagaaatttcggcagcgtgcagttcgtggccagctacaaggtggcct acagcaacgatagcgccaactggaccgagtatcaggaccctagaaccggcagctcca agatcttccccggcaactgggacaaccacagccacaagaagaatctgttcgagacacc catcctggccagatacgtgcggattctgcctgtggcctggcacaacagaatcgccctga gactggaactgctgggctgt
139	EGF_HSA[A626-L633]	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKDAHKSEVAHRFKDLGEENFKALVLIAFAQY LQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQH KDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIA RRHPYFYAPELLFFAKRYKAAFTECCQAADKAACLLPK LDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR LSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADD

		RADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEN DEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY EYARRHPDYSVVLLLRLAKTYETTLEKCCAAADPHEC YAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNAL LVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAK RMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESL VNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKE RQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEK CCKADDKETCFAEEGKKLVA
140	Axit nucleic của 139	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaggatgcccacaagagcgaggtggcccacagattcaag gatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtatctcca gcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcgccaa gacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccctgttc ggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggccgac tgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaaggacgac aaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcctttca cgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggcaccc ctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcaccgaa tgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctgagaga tgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccagaagtt cggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcctaag gccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacaccgagt gctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagtacat ctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaagcccc tgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccgatctg cctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgccgagg ccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccgactattc tgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatgctgcg ccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctctggtgg

		aagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctgggcgagt acaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgtccaca cctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttgcaagc accctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctgaatca gctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgctgtacc gagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagacatacg tgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccctgtcc gagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagcacaa gcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcctttgtc gagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggcaagaa actggtggct
141	Liên kết PS MFG-E8 C1	CVEPLGLENGNIANSQIAASSVRVTFLGLQHWVPELAR LNRAGMVNAWTPSSNDDNPWIQVNLLRRMWVTGVVT QGASRLASHEYLKAFKVAYSLNGHEFDFIHDVNKKHK EFVGNWNKNAVHVNLFETPVEAQYVRLYPTSCHTACT LRFELLGCELNG
142	Liên kết PS MFG-E8 [L76M] C1	CVEPLGMENGNIANSQIAASSVRVTFLGLQHWVPELAR LNRAGMVNAWTPSSNDDNPWIQVNLLRRMWVTGVVT QGASRLASHEYLKAFKVAYSLNGHEFDFIHDVNKKHK EFVGNWNKNAVHVNLFETPVEAQYVRLYPTSCHTACT LRFELLGCELNG
143	Liên kết PS MFG-E8 C2	CANPLGLKNNSIPDKQITASSSYKTWGLHLFSWNPSYA RLDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVTGIIT QGARNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSS KIFPGNWDNHSHKKNLFETPILARYVRILPVAWHNRIA LRLELLGC
144	Liên kết PS EDIL-3 C1	CSGPLGIEGGIISNQQITASSTHRALFGLQKWYPYYARL NKKGLINAWTAAENDRWPWIQINLQRKMRVTGVITQG AKRIGSPEYIKSYKIAYSNDGKTWAMYKVKGTNEDMV

		FRGNIDNNTPYANSFTPPIKAQYVRLYPQVCRRHCTLR MELLGCELSG
145	Liên kết PS EDIL-3 C2	CSEPLGMKSGHIQDYQITASSIFRTLNMDMFTWEPRKA RLDKQGKVNAWTSGHNDQSQWLQVDLLVPTKVTGIIT QGAKDFGHVQFVGSYKLAYSNDGEHWTVYQDEKQRK DKVFQGNFDNDTHRKNVIDPPIYARHIRILPWSWYGRIT LRSELLGCTEEE
146	Liên kết PS EDIL-3 C2 bị cắt cụt TEEE	CSEPLGMKSGHIQDYQITASSIFRTLNMDMFTWEPRKA RLDKQGKVNAWTSGHNDQSQWLQVDLLVPTKVTGIIT QGAKDFGHVQFVGSYKLAYSNDGEHWTVYQDEKQRK DKVFQGNFDNDTHRKNVIDPPIYARHIRILPWSWYGRIT LRSELLGC
147 =117	Trình tự protein FP133	LDICSKNPCHNGGLCEEISQEVRGDVFPSYTCTCLKGYA GNHCETKGSDAHKSEVAHRFKDLGEENFKALVLIAFA QYLQQSPFEDHVKLVNEVTEFAKTCVADESAENCDKS LHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECF LQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYL YEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAC LLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAW AVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLE CADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIA EVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLG MFLYEYARRHPDYSVVLLLRLAKTYETTLEKCCAAAD PHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKH PEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCC TESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAA FVEKCCKADDKETCFAEEGKKLVAASQAALGLGGSGG SGGSGGSCVEPLGMENGNIANSQIAASSVRVTFLGLQH

		WVPELARLNRAGMVNAWTPSSNDDNPWIQVNLLRRM WVTGVVTQGASRLASHEYLKAFKVAYSLNGHEFDFIH DVNKKHKEFVGNWNKNAVHVNLFETPVEAQYVRLYP TSCHTACTLRFELLGCELNG
148	Trình tự axit nucleic FP133	ctggacatctgtagcaagaacccttgccacaacggcggcctgtgcgaagagatttctca agaagtgcggggcgacgttttccccagctacacctgtacatgtctgaagggctacgccg gcaaccactgcgagacaaaaggatctgatgcccacaagagcgaggtggcccacagat tcaaggatctgggcgaagagaacttcaaggccctggtgctgatcgccttcgctcagtat ctccagcagagccctttcgaggaccacgtgaagctggtcaacgaagtgaccgagttcg ccaagacctgtgtggccgatgagagcgccgagaactgtgataagagcctgcacaccct gttcggcgacaagctgtgtacagtggccacactgagagaaacctacggcgagatggc cgactgctgtgccaagcaagagcccgagagaaacgagtgcttcctccagcacaagga cgacaaccccaacctgcctagactcgtgcgacccgaagtggatgtgatgtgcaccgcc tttcacgacaacgaggaaaccttcctgaagaagtacctgtacgagatcgccagacggc acccctacttttatgcccctgagctgctgttcttcgccaagcggtataaggccgccttcac cgaatgttgccaggccgctgataaggctgcctgtctgctgcctaagctggacgagctga gagatgagggcaaagccagctctgccaagcagagactgaaatgcgccagcctccag aagttcggcgagagagcttttaaggcctgggccgttgccagactgagccagagatttcc taaggccgagtttgccgaggtgtccaagctcgtgaccgatctgacaaaggtgcacacc gagtgctgtcacggcgatctgctggaatgtgccgacgatagagccgacctggccaagt acatctgcgagaaccaggacagcatcagcagcaagctgaaagagtgctgcgagaag cccctgctggaaaagtctcactgtatcgccgaggtggaaaacgacgagatgcctgccg atctgcctagcctggctgccgatttcgtggaaagcaaggacgtgtgcaagaactacgcc gaggccaaggatgtgtttctgggcatgtttctgtatgagtacgcccgcagacaccccga ctattctgtggttctgctgctgcggctggccaaaacctacgagacaaccctggaaaaatg ctgcgccgctgccgatcctcacgagtgttatgccaaggtgttcgacgagttcaagcctct ggtggaagaaccccagaacctgatcaagcagaactgcgagctgttcgagcagctggg cgagtacaagttccagaatgccctgctcgtgcggtacaccaagaaagtgcctcaggtgt ccacacctacactggttgaggtgtcccggaatctgggcaaagtgggcagcaagtgttg caagcaccctgaggccaagagaatgccttgcgccgaggattacctgagcgtggtgctg aatcagctgtgcgtgctgcacgagaaaaccccctgtgtccgacagagtgaccaagtgct

		gtaccgagagcctcgtgaacagaaggccttgctttagcgccctggaagtggacgagac atacgtgcccaaagagttcaacgccgagacattcaccttccacgccgacatctgcaccc tgtccgagaaagagcggcagatcaagaagcagacagccctggtcgagctggttaagc acaagcccaaggccaccaaagaacagctgaaggccgtgatggacgacttcgccgcct ttgtcgagaagtgctgcaaggccgacgacaaagagacatgcttcgccgaagagggca agaaactggtggctgcctctcaggctgctctcggacttggtggaagcggaggaagtgg tggatctggcggatcttgtgtggaacccctcggcatggaaaacggcaatatcgccaata gccagattgccgccagcagcgtcagagtgacatttctgggactgcaacactgggtgcc cgagctggctagactgaatagagccggcatggtcaacgcctggacacccagcagcaa cgacgataatccctggattcaagtgaacctgctgcggcgtatgtgggtcacaggtgttgt tacacagggcgcaagcagactggccagccacgagtatctgaaggcctttaaggtggc ctacagcctgaacggccacgagttcgacttcatccacgacgtgaacaagaagcacaaa gagtttgtcggcaactggaacaagaacgccgtgcacgtgaacctgttcgagacacctgt ggaagcccagtacgtgcggctgtaccctacaagctgtcacaccgcctgcactctgagat tcgaactgctgggatgcgagctgaacggc

Sáng chế cũng bao gồm các biến thể của mỗi trình tự trong số các SEQ ID NO: 69, 70 và 72, trong đó miền giống EGF của trình tự EDIL3 được bao gồm ở đó tương ứng với trình tự bất kỳ trong số các trình tự sau đây: SEQ ID NO: 96, SEQ ID NO: 97, SEQ ID NO: 98, SEQ ID NO: 99, SEQ ID NO: 100, hoặc SEQ ID NO: 101.

Sáng chế cũng bao gồm protein dung hợp trị liệu bao gồm các miền liên kết integrin của MFGE8 hoặc EDIL3, và các miền liên kết PS như miền IgSF V của TIM4 hoặc miền GLA của các biến thể GAS6 của protein bắc cầu (ví dụ FP1147 và FP1148).

Cải biến của các protein theo sáng chế

Sáng chế cũng bao gồm các biến thể của các protein được mô tả ở đây và/hoặc các mảnh của chúng có các cải biến khác nhau ở các miền cũng như các thể dung hợp và thể liên hợp của các phân tử được bộc lộ. Ví dụ, miền của protein dung hợp trị liệu có thể có cải biến bảo toàn của các gốc axit amin, và trong đó các protein đã cải biến vẫn giữ hoặc có các tính chất được tăng cường như được so sánh với protein dung hợp bao gồm miền ban đầu. Theo cách khác, miền của protein dung hợp trị liệu có thể có (các) sự khuyết gốc axit amin, trong đó các protein dung hợp được cải biến vẫn giữ hoặc có các tính chất được tăng cường so với protein bao gồm miền ban đầu. Theo cách khác,

các protein dung hợp trị liệu có thể có (các) sự chèn gốc axit amin, trong đó các protein được cải biến này vẫn giữ hoặc có các tính chất được tăng cường so với protein không được cải biến. Theo một phương án, sự chèn axit amin như vậy bao gồm các gốc glyxin hoặc serin trong nhiều tổ hợp để hoạt động là cầu nối giữa các miền của protein ban đầu.

Sự gây đột biến định hướng vị trí hoặc gây đột biến được trung gian bởi PCR có thể được thực hiện để đưa vào (các) đột biến và tác dụng trên liên kết integrin và/hoặc PS, hoặc tính chất chức năng quan tâm khác, có thể được đánh giá trong các thử nghiệm *in vitro* hoặc *in vivo*. Các cải biến bảo toàn (như được thảo luận ở trên) có thể được đưa vào và/hoặc các đột biến có thể là sự thế, sự bổ sung hoặc sự loại bỏ axit amin. Hơn thế nữa, thông thường không có nhiều hơn một, hai, ba, bốn hoặc năm gốc bên trong miền liên kết được thay đổi.

Các biến thể trình tự axit amin của các protein dung hợp trị liệu, mà có các tính chất về cơ bản tương tự là các biến thể không được cải biến, có thể được tạo bằng cách đưa các sự thay đổi nucleotit thích hợp vào trong các ADN mã hóa, hoặc bằng sự tổng hợp các biến thể mong muốn. Các biến thể như vậy bao gồm, ví dụ, sự loại bỏ khỏi, hoặc sự chèn hoặc sự thế của, các gốc bên trong các trình tự axit amin của các phân tử này. Theo một số phương án, các biến thể có thể bao gồm các trình tự cầu nối bổ sung, các trình tự cầu nối được giảm hoặc sự loại bỏ các trình tự cầu nối, và/hoặc sự đột biến hoặc sự thế axit amin và sự loại bỏ một hoặc nhiều axit amin. Sự kết hợp bất kỳ của sự loại bỏ, sự chèn và sự thế được tạo ra để đến cấu trúc cuối cùng, miễn là cấu trúc cuối cùng này có các đặc tính mong muốn. Sự thay đổi axit amin này cũng có thể thay đổi các quy trình sau dịch mã của phân tử, như thay đổi số lượng hoặc vị trí của các vị trí glycosyl hóa có thể có.

Phương pháp tạo ra phân tử tái tổ hợp

Axit nucleic và hệ biểu hiện

Theo một phương án, sáng chế đề xuất phương pháp tạo ra một hoặc nhiều chuỗi polypeptit của protein dung hợp trị liệu theo cách tái tổ hợp, bao gồm bước: 1) sản xuất một hoặc nhiều cấu trúc ADN chứa phân tử axit nucleic mã hóa chuỗi polypeptit của phân tử liên kết đa đặc hiệu; 2) đưa (các) cấu trúc ADN này vào trong một hoặc nhiều vectơ biểu hiện; 3) đồng chuyển nạp (các) vectơ biểu hiện này trong một hoặc nhiều tế bào chủ; và 4) biểu hiện và lắp ráp phân tử này trong tế bào chủ hoặc trong dung dịch.

Trong khía cạnh này, sáng chế đề xuất các axit nucleic được phân lập, ví dụ, một hoặc nhiều polynucleotit, mã hóa các protein dung hợp trị liệu được mô tả ở đây. Các phân tử axit nucleic bao gồm ADN và ARN trong cả dạng sợi đơn và sợi kép, cũng như các trình tự bổ sung tương ứng. Các phân tử axit nucleic theo sáng chế bao gồm các gen chiều dài đầy đủ hoặc các phân tử cADN cũng như sự kết hợp của các mảnh của chúng. Các axit nucleic theo sáng chế thu được từ các nguồn người nhưng sáng chế bao gồm các loại thu được từ các loài không phải là người.

'Axit nucleic được phân lập' là axit nucleic mà đã được tách ra từ các trình tự di truyền liền kề có mặt trong bộ gen của sinh vật mà axit nucleic được phân lập từ đó, trong trường hợp các axit nucleic được phân lập từ các nguồn xuất hiện tự nhiên. Trong trường hợp axit nucleic được tổng hợp bằng enzym từ khuôn hoặc tổng hợp hóa học, như sản phẩm PCR, phân tử cADN, hoặc oligonucleotit ví dụ, cần hiểu rằng các axit nucleic thu được từ các quy trình như vậy là các axit nucleic được phân lập. Phân tử axit nucleic được phân lập là để chỉ phân tử axit nucleic ở dạng đoạn riêng biệt hoặc là thành phần của cấu trúc axit nucleic lớn hơn. Trong một phương án được ưu tiên, các axit nucleic về cơ bản không nhiễm bẩn vật liệu nội sinh. Các phân tử axit nucleic tốt hơn là thu được từ ADN hoặc ARN được phân lập ít nhất một lần ở dạng về cơ bản tinh khiết và ở lượng hoặc nồng độ cho phép xác định, thao tác, và thu hồi các trình tự nucleotit thành phần của nó bằng các phương pháp hóa sinh chuẩn (như các phương pháp được nêu trong tài liệu Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989)). Tốt hơn là, các trình tự như vậy được tạo ra và/hoặc được cấu trúc ở dạng khung đọc mở không bị gián đoạn bởi các nội trình tự không được dịch mã, hoặc các intron, thường có mặt trong các gen của tế bào nhân chuẩn. Các trình tự của ADN không được dịch mã có thể có mặt ở đầu 5' hoặc 3' từ khung đọc mở, trong đó nó không ảnh hưởng đến sự thao tác hoặc biểu hiện của vùng mã hóa.

Sáng chế cũng đề xuất hệ biểu hiện và cấu trúc ở dạng plasmit, vectơ biểu hiện, catxet dịch mã hoặc biểu hiện, mà bao gồm ít nhất một polynucleotit như được mô tả trên đây. Ngoài ra, sáng chế đề xuất tế bào chủ chứa hệ hoặc cấu trúc biểu hiện này.

Theo một phương án, sáng chế đề xuất phương pháp tạo ra protein dung hợp trị liệu bao gồm các bước: (a) nuôi cấy tế bào chủ chứa axit nucleic mã hóa protein dung

hợp, trong đó tế bào chủ được nuôi cấy biểu hiện protein dung hợp; và (b) thu hồi protein dung hợp từ dịch nuôi cấy tế bào chủ.

Cũng được đề xuất trong sáng chế là vectơ biểu hiện và tế bào chủ sản xuất các protein dung hợp trị liệu được mô tả trên đây. Thuật ngữ "vectơ" nghĩa là phân tử hoặc thực thể bất kỳ (ví dụ axit nucleic, plasmit, thể thực khuẩn hoặc virut) mà thích hợp cho sự biến nạp hoặc chuyển nhiễm của tế bào chủ và chứa các trình tự axit nucleic mà điều khiển và/hoặc kiểm soát (trong sự liên hợp với tế bào chủ) sự biểu hiện của một hoặc nhiều vùng mã hóa khác loại được gắn hoạt động với chúng. Các vectơ biểu hiện khác nhau có thể được sử dụng để biểu hiện các chuỗi mã hóa các polynucleotit hoặc các miền gắn của phân tử. Cả các vectơ biểu hiện dựa trên virut và không phải virut đều có thể được sử dụng để tạo ra protein dung hợp trị liệu trong tế bào vật chủ của động vật có vú. Các vectơ và hệ thống không phải virut bao gồm các plasmit, các vectơ episom, điển hình có catxet biểu hiện để biểu hiện protein hoặc ARN, và các nhiễm sắc thể nhân tạo của con người (xem, ví dụ, Harrington *et al.*, (1997) Nat Genet 15: 345). Ví dụ, các vectơ không phải virut hữu ích để biểu hiện các polynucleotit và các polypeptit trong tế bào động vật có vú (ví dụ, người) bao gồm pThioHis A, B & C, pcDNA3.1 / His, pEBVHis A, B & C, (Invitrogen, San Diego, CA), các vectơ MPSV, và nhiều vectơ khác được biết đến trong lĩnh vực kỹ thuật để biểu hiện các protein khác. Các vectơ virut hữu ích bao gồm các vectơ dựa trên các retrovirut, các adenovirut, các virut liên kết adeno, các virut herpe, các vectơ dựa trên SV40, virut u nhú, virut HBP Epstein Barr, các vectơ virut vaccin và virut Semliki Forest (SFV). Xem, Brent *et al.*, (1995) nêu trên; Smith, Annu. Rev. Microbiol. 49: 807; và Rosenfeld et al., (1992) Cell 68: 143.

Việc lựa chọn vectơ biểu hiện phụ thuộc vào các tế bào chủ dự định trong đó vectơ sẽ được biểu thị. Thông thường, các vectơ biểu hiện chứa vùng khởi động và các trình tự điều hòa khác (ví dụ, các vùng tăng cường) được liên kết hoạt động với các polynucleotit mã hóa protein dung hợp trị liệu. Theo một số phương án, vùng khởi động cảm ứng được sử dụng để ngăn chặn sự biểu hiện của các trình tự được chèn vào ngoại trừ trong các điều kiện gây cảm ứng. Vùng khởi động cảm ứng bao gồm, ví dụ, vùng khởi động arabinoza, lacZ, metallothionein hoặc vùng khởi động sốc nhiệt. Việc nuôi cấy sinh vật biến nạp có thể được mở rộng trong điều kiện không cảm ứng mà không thiên về quần thể đối với các trình tự mã hóa mà sản phẩm biểu hiện của chúng được dung nạp tốt hơn bởi tế bào chủ. Ngoài các vùng khởi động, các yếu tố điều hòa khác

cũng có thể được yêu cầu hoặc mong muốn để biểu hiện hiệu quả các protein dung hợp trị liệu. Các thành phần này thường bao gồm codon mở đầu ATG và vị trí liên kết ribosom liền kề hoặc trình tự khác. Ngoài ra, hiệu quả biểu hiện có thể được tăng cường bằng cách bao gồm các vùng tăng cường thích hợp với hệ thống tế bào đang sử dụng (xem, ví dụ, Scharf *et al.*, (1994) Results Probl. Cell Differ. 20: 125; và Bittner et al., (1987) Meth. Enzymol., 153 :516). Ví dụ như, vùng tăng cường SV40 hoặc vùng tăng cường CMV có thể được sử dụng để làm tăng sự biểu hiện trong tế bào chủ động vật có vú.

Các vectơ biểu hiện cũng có thể cung cấp vị trí trình tự tín hiệu tiết để tạo thành protein dung hợp với các polypeptit được mã hóa bằng cách chèn các trình tự được mô tả ở trên của các miền gắn và/hoặc các miền hòa tan. Thông thường hơn, các trình tự được chèn được liên kết với các trình tự tín hiệu trước khi đưa vào vectơ. Các vectơ cho phép biểu hiện các miền gắn và miền hòa tan dưới dạng các protein dung hợp do đó dẫn đến việc sản xuất các protein được thao tác di truyền nguyên vẹn. Tế bào chủ, khi được nuôi cấy trong các điều kiện thích hợp, có thể được sử dụng để biểu hiện protein được thao tác di truyền mà sau đó có thể được thu thập từ môi trường nuôi cấy (nếu tế bào chủ tiết ra nó vào môi trường) hoặc trực tiếp từ tế bào chủ sản xuất ra nó (nếu không được tiết ra). Việc lựa chọn tế bào chủ thích hợp sẽ phụ thuộc vào nhiều yếu tố khác nhau, như các mức độ biểu hiện mong muốn, các cải biến polypeptit mong muốn hoặc cần thiết cho hoạt động (như glycosyl hóa hoặc phosphoryl hóa) và dễ dàng gấp lại thành phân tử có hoạt tính sinh học. Tế bào chủ có thể là tế bào nhân chuẩn hoặc tế bào nhân sơ.

Các dòng tế bào động vật có vú có sẵn làm các vật chủ để biểu hiện đã được biết đến trong lĩnh vực và bao gồm, nhưng không giới hạn, các dòng tế bào bất tử có sẵn từ Bộ sưu tập chủng chuẩn Hoa Kỳ (ATCC) và bất kỳ dòng tế bào nào được sử dụng trong hệ thống biểu hiện được biết đến trong lĩnh vực có thể được sử dụng để tạo ra các protein dung hợp tái tổ hợp theo sáng chế. Nói chung, các tế bào chủ được biến nạp với vectơ biểu hiện tái tổ hợp bao gồm ADN mã hóa protein dung hợp mong muốn. Trong số các tế bào chủ có thể được sử dụng là các tế bào nhân sơ, các nấm men hoặc các tế bào nhân thực cấp cao hơn. Tế bào nhân sơ bao gồm tế bào từ sinh vật gram âm hoặc gram dương, ví dụ E. *coli* hoặc bacilli. Tế bào nhân chuẩn bậc cao hơn bao gồm tế bào côn trùng và các dòng tế bào được tạo ra có nguồn gốc từ động vật có vú. Các ví dụ của các dòng tế

bào vật chủ thích hợp của động vật có vú bao gồm tế bào COS-7, tế bào L, tế bào Cl27, tế bào 3T3, tế bào buồng trứng hamster Trung Quốc (CHO) hoặc các dẫn xuất của chúng và các dòng tế bào liên quan phát triển trong môi trường tự do huyết thanh, tế bào HeLa, tế bào BHK dòng, dòng tế bào CV-1 EBNA, tế bào thận phôi người (HEK) như 293, 293 EBNA hoặc MSR 293, tế bào A431 biểu bì của người, tế bào Colo205 ở người, các dòng tế bào linh trưởng đã biến đổi khác, tế bào lưỡng bội bình thường, chủng tế bào có nguồn gốc từ nuôi cấy in vitro mô sơ cấp, mẫu cấy sơ cấp, tế bào HL-60, U937, HaK hoặc Jurkat. Tùy ý, các dòng tế bào của động vật có vú như HepG2/3B, KB, NIH 3T3 hoặc S49, ví dụ, có thể được sử dụng để biểu hiện polypeptit nếu muốn sử dụng polypeptit này trong các thử nghiệm sàng lọc hoặc truyền tín hiệu khác nhau. Theo cách khác, có thể tổng hợp polypeptit trong tế bào nhân chuẩn bậc thấp như tế bào nấm men hoặc trong tế bào nhân sơ như vi khuẩn. Các loại nấm men thích hợp bao gồm các chủng *P. pastoris*, *S. cerevisiae, S. pombe*, *Kluyveromyces*, *Candida*, hoặc bất kỳ chủng nấm men nào có khả năng biểu hiện các polypeptit dị hợp. Các chủng vi khuẩn thích hợp bao gồm *E. coli, B. subtilis, S. typhimurium*, hoặc bất kỳ chủng vi khuẩn nào có khả năng biểu hiện các polypeptit dị hợp. Nếu protein dung hợp được tạo ra từ nấm men hoặc vi khuẩn, thì có thể mong muốn cải biến sản phẩm được tạo ra từ đó, ví dụ bằng cách phosphoryl hóa hoặc glycosyl hóa ở các vị trí thích hợp, để thu được sản phẩm chức năng. Sự liên kết cộng hóa trị như vậy có thể được thực hiện bằng cách sử dụng các phương pháp hóa học hoặc enzym đã biết.

Phương pháp để đưa vectơ biểu hiện chứa trình tự polynucleotit được quan tâm thay đổi tùy thuộc vào kiểu vật chủ tế bào. Ví dụ, chuyển nhiễm canxi clorua thường được sử dụng cho tế bào sinh vật nhân sơ, trong khi đó xử lý canxi phosphat hoặc điện di có thể được sử dụng cho các vật chủ tế bào khác. Các phương pháp khác bao gồm, ví dụ, xung điện, xử lý canxi phosphat, biến nạp qua trung gian liposom, tiêm và vi tiêm, phương pháp bắn gen, virosom, immunoliposom, thể liên hợp polycation:axit nucleic, ADN trần, virion nhân tạo, dung hợp với protein cấu trúc virut herpes VP22, sự hấp thụ ADN được tăng cường bằng chất, và sự tải nạp *ex vivo*. Để sản xuất năng suất cao, trong thời gian lâu dài các protein tái tổ hợp, sự biểu hiện ổn định thường được mong muốn. Ví dụ, các dòng tế bào biểu hiện ổn định các protein được thao tác di truyền có thể được chuẩn bị bằng cách sử dụng vectơ biểu hiện theo sáng chế chứa nguồn gốc virut của các yếu tố sao chép hoặc biểu hiện nội sinh và gen đánh dấu có thể lựa chọn. Sau khi đưa

vectơ vào, tế bào có thể được để cho sinh trưởng trong 1-2 ngày trong môi trường được làm giàu trước khi chúng được chuyển sang môi trường chọn lọc. Mục đích của gen đánh dấu chọn lọc là mang lại tính kháng để chọn lọc, và sự có mặt của nó cho phép sự sinh trưởng của tế bào mà biểu hiện thành công trình tự đã được đưa vào trong môi trường chọn lọc. Tế bào được chuyển nạp ổn định, có tính kháng có thể được tăng sinh bằng cách sử dụng kỹ thuật nuôi cấy mô thích hợp với loại tế bào.

Các protein dung hợp thường được phục hồi từ môi trường nuôi cấy dưới dạng polypeptit được tiết ra, mặc dù chúng cũng có thể được khôi phục từ dịch phân ly của tế bào chủ khi được sản xuất trực tiếp mà không có tín hiệu bài tiết. Nếu polypeptit được liên kết với màng, nó có thể được giải phóng khỏi màng bằng cách sử dụng dung dịch tẩy rửa thích hợp (ví dụ, Triton-X 100).

Khi protein dung hợp được tạo ra trong tế bào tái tổ hợp không phải tế bào có nguồn gốc từ người, thì nó hoàn toàn không chứa protein hoặc polypeptit có nguồn gốc từ người. Tuy nhiên, cần phải tinh sạch protein dung hợp từ các protein tế bào tái tổ hợp hoặc các polypeptit. Bước đầu tiên, môi trường nuôi cấy hoặc dịch lọc thường được ly tâm để loại bỏ các mảnh vụn tế bào dạng hạt. Các phân tử được tạo ra có thể được tinh chế một cách thuận tiện bằng sắc ký hydroxylapatit, điện di gel, thẩm tách, hoặc sắc ký ái lực, với sắc ký ái lực là kỹ thuật tinh sạch được ưu tiên. Các kỹ thuật khác để tinh sạch protein như phân đoạn trên cột trao đổi ion, kết tủa bằng etanol, pha ngược HPLC, sắc ký trên silic oxit, sắc ký trên heparin Sepharose, sắc ký trên nhựa trao đổi anion hoặc cation (như cột axit đa chức), sắc ký, SDS-PAGE, và kết tủa amoni sulfat cũng có sẵn.

Theo một số khía cạnh nhất định, được đề xuất ở đây là vectơ virut bao gồm polynucleotit mã hóa protein dung hợp trị liệu theo sáng chế. Trong một số phương án, vectơ virut có nguồn gốc từ AAV. Theo một số phương án nhất định, vectơ virut được dùng cho đối tượng, ví dụ, con người, trong đó protein dung hợp trị liệu được biểu hiện và có thể được sử dụng để điều trị và/hoặc phòng ngừa các bệnh như được liệt kê ở đây.

Dược phẩm

Theo một khía cạnh khác, sáng chế đề xuất chế phẩm, ví dụ, dược phẩm, chứa protein dung hợp trị liệu theo sáng chế, kết hợp với một hoặc nhiều tá dược, chất pha loãng hoặc chất mang dược dụng. Các chế phẩm như vậy có thể bao gồm một hoặc sự kết hợp của (ví dụ, hai hoặc nhiều) protein dung hợp trị liệu khác nhau theo sáng chế.

Các dược phẩm như được mô tả ở đây cũng có thể được dùng trong liệu pháp phối hợp, tức là, được kết hợp với các thuốc khác. Ví dụ, liệu pháp kết hợp có thể bao gồm protein dung hợp theo sáng chế được kết hợp với, ví dụ, ít nhất một chất chống viêm, chống nhiễm trùng hoặc chất ức chế miễn dịch. Các ví dụ của các chất trị liệu có thể được sử dụng trong liệu pháp kết hợp được mô tả chi tiết hơn bên dưới trong phần về việc sử dụng các protein dung hợp trị liệu theo sáng chế.

Để điều chế dược phẩm hoặc chế phẩm vô trùng bao gồm protein dung hợp theo sáng chế, protein dung hợp được trộn với chất mang hoặc tá dược dược dụng.

Cụm từ 'dược dụng' có nghĩa là được cơ quan quản lý của chính quyền liên bang hoặc tiểu bang chấp thuận hoặc được liệt kê trong Dược điển Hoa Kỳ hoặc dược lý được công nhận chung khác để sử dụng trên động vật và đặc biệt hơn là ở người.

Thuật ngữ 'dược phẩm' dùng để chỉ hỗn hợp của ít nhất một thành phần hoạt tính (ví dụ, protein được thiết kế) và ít nhất một tá dược dược dụng, chất pha loãng hoặc chất mang.

'Thuốc' để chỉ một chất được sử dụng để điều trị y khoa.

Như được sử dụng ở đây, 'chất mang dược dụng' bao gồm bất kỳ và tất cả dung môi, môi trường phân tán, chất phủ, chất kháng khuẩn và chống nấm, chất đẳng trương và chất làm chậm hấp thu, và những chất tương tự tương thích về mặt sinh lý. Chất mang phải phù hợp để tiêm tĩnh mạch, tiêm bắp, tiêm dưới da, tiêm, tủy sống hoặc biểu bì (ví dụ, bằng cách tiêm hoặc truyền). Theo một phương án, chất mang phải phù hợp với đường tiêm dưới da. Tùy thuộc vào cách sử dụng, hợp chất hoạt động, tức là protein dung hợp, có thể được phủ trong một vật liệu để bảo vệ hợp chất khỏi tác dụng của axit và các điều kiện tự nhiên khác có thể làm mất hoạt tính của hợp chất.

Dược phẩm như được mô tả ở đây có thể bao gồm một hoặc nhiều muối dược dụng. Dược phẩm như được mô tả ở đây cũng có thể bao gồm chất chống oxi hóa dược dụng. Ví dụ về các chất chống oxi hóa dược dụng bao gồm: chất chống oxi hóa hòa tan trong nước, như axit ascorbic, xystein hydroclorua, natri bisulfit, natri metabisulfit, natri sulfit và các chất tương tự; chất chống oxi hóa tan trong dầu, ascorbyl palmitat, hydroxyanisol được butyl hóa (BHA), hydroxytoluen được butyl hóa (BHT), lecithin, propyl gallat, alpha-tocopherol, và các loại tương tự; và các chất chelat hóa kim loại,

như axit xitric, axit etylendiamin tetraaxetic (EDTA), sorbitol, axit tartaric, axit phosphoric, và các loại tương tự.

Các ví dụ của các chất mang nước và không phải nước thích hợp có thể được sử dụng trong dược phẩm như được mô tả ở đây bao gồm nước, etanol, polyol (như glyxerol, propylen glycol, polyetylen glycol, và các chất tương tự), và các hỗn hợp thích hợp của chúng, dầu thực vật, như dầu ô liu, và các este hữu cơ dạng tiêm, như etyl oleat. Độ chảy thích hợp có thể được duy trì, ví dụ, bằng cách sử dụng các vật liệu phủ, như lexithin, bằng cách duy trì của kích cỡ hạt cần thiết trong trường hợp các thể phân tán, và bằng cách sử dụng các chất hoạt động bề mặt.

Các chế phẩm này cũng có thể chứa các chất bổ trợ như chất bảo quản, chất làm ướt, chất nhũ hóa và chất phân tán. Có thể đảm bảo ngăn ngừa sự hiện diện của vi sinh vật bằng cả quy trình khử trùng và bằng cách bao gồm các chất kháng khuẩn và kháng nấm khác nhau, ví dụ, paraben, chlorobutanol, axit phenol sorbic, và các chất tương tự. Cũng có thể mong muốn bao gồm các chất đẳng trương, như đường, natri clorua, và các chất tương tự vào chế phẩm. Ngoài ra, sự hấp thu kéo dài của dạng dược phẩm tiêm có thể được tạo ra bởi việc bao gồm các tác nhân làm chậm hấp thu như nhôm monostearat và gelatin.

Chất mang dược dụng bao gồm thể phân tán hoặc dung dịch dạng nước vô trùng và bột vô trùng để tạo ra thể phân tán hoặc dung dịch vô trùng có thể tiêm được ngay lúc dùng. Việc sử dụng các phương tiện và chất như vậy cho các chất dược dụng đã được biết đến trong lĩnh vực này. Ngoại trừ trong trường hợp mà môi trường hoặc chất thông thường bất kỳ không tương thích với hợp chất hoạt tính, việc sử dụng nó trong dược phẩm theo sáng chế được dự tính. Hợp chất có hoạt tính bổ sung cũng có thể được kết hợp vào các chế phẩm.

Các chế phẩm trị liệu thường phải vô trùng và ổn định trong các điều kiện sản xuất và bảo quản. Dược phẩm này có thể được bào chế dưới dạng dung dịch, vi nhũ tương, liposom hoặc các cấu trúc được sắp xếp khác thích hợp với nồng độ thuốc cao. Chất mang có thể là dung môi hoặc môi trường phân tán chứa, ví dụ, nước, etanol, rượu đa chức (ví dụ, glyxerol, propylen glycol, và polyetylen glycol dạng lỏng, và chất tương tự), và hỗn hợp thích hợp của chúng. Có thể duy trì độ lỏng phù hợp, ví dụ, bằng cách sử dụng chất phủ như lexithin, bằng cách duy trì kích thước hạt yêu cầu trong trường

hợp của dịch phân tán và bằng cách sử dụng chất hoạt động bề mặt. Trong nhiều trường hợp, có thể bao gồm các chất đẳng trương, ví dụ, đường, các rượu đa chức như mannitol, sorbitol, hoặc natri clorua trong chế phẩm.

Các đánh giá về sự phát triển của các công thức protein ổn định có thể được tìm thấy trong Cleland *et al.,* (1993) Crit Reviews Ther Drug Carrier Systems, 10(4): 307-377 và Wei W (1999) Int J Pharmaceutics, 185: 129-88.

Dung dịch hoặc huyền phù dùng để bôi trong da hoặc dưới da thường bao gồm một hoặc nhiều thành phần sau: chất pha loãng vô trùng như nước pha tiêm, dung dịch muối, dầu cố định, polyetylen glycol, glyxerin, propylen glycol hoặc dung môi tổng hợp khác, chất kháng khuẩn như rượu benzyl hoặc các metyl paraben, chất chống oxi hóa như axit ascorbic hoặc natri bisulfit, chất tạo chelat như axit etylendiaminotetraaxetic, chất đệm như axetat, xitrat hoặc phosphat và các chất điều chỉnh trương lực như natri clorua hoặc dextroza. Độ pH có thể được điều chỉnh bằng axit hoặc bazơ, như axit clohydric hoặc natri hydroxit. Các chế phẩm như vậy có thể được đựng trong ống tiêm, ống tiêm dùng một lần hoặc nhiều lọ đựng liều lượng bằng thủy tinh hoặc nhựa.

Các dung dịch tiêm vô trùng có thể được điều chế bằng cách cho hợp chất hoạt động với lượng cần thiết vào dung môi thích hợp với một hoặc kết hợp các thành phần được liệt kê ở trên, theo yêu cầu, sau đó là vi lọc khử trùng. Nói chung, các chất phân tán được điều chế bằng cách kết hợp các protein dung hợp theo sáng chế vào phương tiện vô trùng có chứa môi trường phân tán cơ bản và các thành phần khác cần thiết từ các thành phần được liệt kê ở trên. Trong trường hợp bột vô trùng để điều chế dung dịch tiêm vô trùng, các phương pháp điều chế là sấy khô chân không và sấy khô đông lạnh (đông khô) để tạo ra bột của thành phần hoạt tính cộng với bất kỳ thành phần mong muốn bổ sung nào từ dung dịch đã được lọc vô trùng trước đó của nó.

Lượng thành phần hoạt tính có thể được kết hợp với chất mang để tạo ra dạng bào chế đơn sẽ khác nhau tùy thuộc vào đối tượng được điều trị và phương thức sử dụng cụ thể. Lượng hoạt chất có thể được kết hợp với nguyên liệu mang để tạo ra dạng liều dùng đơn nói chung sẽ là lượng chế phẩm tạo ra hiệu quả trị liệu. Nói chung, trong số một trăm phần trăm, lượng này sẽ nằm trong khoảng từ 0,01% đến khoảng chín mươi chín phần trăm thành phần hoạt tính, từ khoảng 0,1 phần trăm đến khoảng 70 phần trăm,

hoặc từ khoảng 1 phần trăm đến khoảng 30 phần trăm thành phần hoạt tính. kết hợp với chất mang dược dụng.

Việc lựa chọn phác đồ sử dụng đối với protein được thao tác di truyền trị liệu phụ thuộc vào một số yếu tố, bao gồm tốc độ luân chuyển mô hoặc huyết thanh của thực thể, mức độ các triệu chứng, khả năng sinh miễn dịch của thực thể, và khả năng tiếp cận của các tế bào đích trong chất nền sinh học. Theo các phương án nhất định, phác đồ sử dụng tối đa hóa hàm lượng của chất trị liệu được chuyển cho bệnh nhân thống nhất với mức độ chấp nhận được của tác dụng phụ. Theo đó, lượng protein được cung cấp phụ thuộc một phần vào thực thể cụ thể và mức độ nghiêm trọng của tình trạng đang được điều trị. Có hướng dẫn lựa chọn các liều lượng sinh học và các phân tử nhỏ thích hợp (xem, ví dụ, Bach (ed.) (1993) Monoclonal Antibodies and Peptide Therapy in Autoimmune Diseases, Marcel Dekker, New York, N.Y.; Baert, et al. (2003) New Engl. J. Med. 348:601-608; Milgrom, et al. (1999) New Engl. J. Med. 341:1966-1973; Slamon, et al. (2001) New Engl. J. Med. 344:783-792; Beniaminovitz, et al. (2000) New Engl. J. Med. 342:613-619; Ghosh, et al. (2003) New Engl. J. Med. 348:24-32; Lipsky, et al. (2000) New Engl. J. Med. 343:1594-1602).

Việc xác định liều lượng thích hợp được thực hiện bởi thầy thuốc lâm sàng, ví dụ, bằng cách sử dụng các thông số hoặc yếu tố đã biết hoặc đã nghi ngờ trong lĩnh vực là ảnh hưởng tới việc điều trị hoặc được dự đoán là ảnh hưởng tới việc điều trị. Thông thường, liều lượng bắt đầu với lượng nhỏ hơn một chút so với liều lượng tối ưu và nó được tăng lên với các lượng gia tăng nhỏ sau đó cho đến khi đạt được tác dụng mong muốn hoặc tối ưu tương quan với tác dụng phụ có hại bất kỳ. Các số đo chẩn đoán quan trọng bao gồm số đo của triệu chứng của, ví dụ, sự viêm hoặc hàm lượng xytokin viêm được sản xuất ra.

Các mức liều dùng thực tế của các thành phần hoạt tính trong dược phẩm theo sáng chế có thể thay đổi để thu được lượng hoạt chất có hiệu quả để đạt được đáp ứng trị iệu mong muốn cho bệnh nhân cụ thể, chế phẩm và cách sử dụng, mà không gây độc cho bệnh nhân. Mức liều dùng được lựa chọn sẽ phụ thuộc vào nhiều yếu tố dược động học bao gồm hoạt tính của các chế phẩm cụ thể theo sáng chế được sử dụng, đường sử dụng, thời gian sử dụng, tốc độ bài tiết của hợp chất cụ thể được sử dụng, thời gian sử dụng điều trị, các loại thuốc, hợp chất và/hoặc vật liệu khác được sử dụng kết hợp với

các chế phẩm cụ thể được sử dụng, tuổi, giới tính, cân nặng, tình trạng, sức khỏe chung và tiền sử bệnh của bệnh nhân đang được điều trị, và các yếu tố đã biết trong y học.

Các phác đồ liều dùng được điều chỉnh để mang lại đáp ứng mong muốn tối ưu. Ví dụ, liều duy nhất có thể được dùng, nhiều liều chia nhỏ có thể được dùng theo thời gian hoặc liều lượng có thể được giảm hoặc tăng tương ứng theo các chỉ dẫn của tình trạng trị liệu. Đặc biệt thuận lợi khi bào chế các chế phẩm dùng đường tiêm dưới dạng đơn vị liều dùng để dễ sử dụng và đồng nhất về liều dùng. Dạng đơn vị liều lượng như được dùng trong bản mô tả này dùng để chỉ các đơn vị riêng rẽ về mặt vật lý thích hợp làm liều lượng đơn nguyên cho đối tượng cần điều trị; mỗi đơn vị chứa lượng định trước của hoạt chất được tính để tạo ra tác dụng trị liệu mong muốn khi kết hợp với chất mang dược cần thiết. Đặc điểm kỹ thuật cho các dạng đơn vị liều dùng theo sáng chế được quy định và phụ thuộc trực tiếp vào các đặc tính riêng biệt của hợp chất hoạt động và hiệu quả trị liệu cụ thể cần đạt được, và những hạn chế vốn có trong lĩnh vực tạo hợp chất hoạt tính như vậy để điều trị nhạy cảm ở cá nhân.

Đối với việc dùng protein dung hợp trị liệu, liều dùng nằm trong khoảng từ 0,0001 đến 150 mg/kg, như 5, 15 và 50 mg/kg dùng dưới da, và thường là từ 0,01 đến 5 mg/kg đối với khối lượng cơ thể vật chủ. Phác đồ điều trị ví dụ bao gồm việc dùng thuốc mỗi tuần một lần, hai tuần một lần, ba tuần một lần, bốn tuần một lần, mỗi tháng một lần, 3 tháng một lần hoặc ba đến 6 tháng một lần.

Các protein dung hợp trị liệu theo sáng chế có thể được dùng nhiều lần. Các quãng cách giữa các liều đơn lẻ có thể là, ví dụ, hàng tuần, hàng tháng, ba tháng một lần hoặc hàng năm. Các quãng cách cũng có thể không đều như được chỉ định bằng cách đo nồng độ protein được thiết kế trong máu ở bệnh nhân. Trong một số phương pháp, liều dùng được điều chỉnh để đạt được nồng độ protein huyết tương khoảng 1-1000 µg/ml và trong một số phương pháp khoảng 25-300 µg/ml.

Ngoài ra, protein dung hợp trị liệu có thể được sử dụng như một công thức giải phóng bền vững, trong trường hợp đó, cần sử dụng ít thường xuyên hơn. Liều dùng và tần suất thay đổi tùy thuộc vào thời gian bán hủy của protein ở bệnh nhân và có thể thay đổi tùy thuộc vào việc liệu việc điều trị là dự phòng hay điều trị. Trong sử dụng phòng ngừa, liều dùng tương đối thấp được sử dụng ở khoảng thời gian tương đối ít thường xuyên trong khoảng thời gian dài. Một số bệnh nhân có thể tiếp tục được điều trị đến hết

đời. Trong các ứng dụng trị liệu, đôi khi cần liều lượng tương đối cao với các khoảng thời gian tương đối ngắn cho đến khi sự tiến triển của tình trạng hoặc bệnh giảm hoặc chấm dứt hoặc cho đến khi bệnh nhân cải thiện một phần hoặc hoàn toàn các triệu chứng của tình trạng hoặc bệnh. Sau đó, bệnh nhân có thể được sử dụng phác đồ phòng bệnh.

Mức liều lượng thực sự của thành phần hoạt tính trong dược phẩm theo sáng chế có thể thay đổi sao cho đạt được lượng của thành phần hoạt tính mà hữu hiệu để đạt được đáp ứng điều trị mong muốn đối với bệnh nhân, dược phẩm và cách thức sử dụng cụ thể, mà không gây độc cho bệnh nhân. Mức liều dùng được lựa chọn sẽ phụ thuộc vào nhiều yếu tố dược động học bao gồm hoạt tính của các chế phẩm cụ thể theo sáng chế được sử dụng, đường sử dụng, thời gian sử dụng, tốc độ bài tiết của hợp chất cụ thể được sử dụng, thời gian sử dụng điều trị, các loại thuốc, hợp chất và/hoặc vật liệu khác được sử dụng kết hợp với các chế phẩm cụ thể được sử dụng, tuổi, giới tính, cân nặng, tình trạng, sức khỏe chung và tiền sử bệnh của bệnh nhân đang được điều trị, và các yếu tố đã biết trong y học.

'Liều dùng hiệu quả trị liệu' của protein dung hợp theo sáng chế có thể làm giảm mức độ nghiêm trọng của tình trạng hoặc các triệu chứng hoặc bệnh và/hoặc ngăn ngừa suy giảm hoặc khuyết tật do tình trạng này.

Chế phẩm theo sáng chế có thể được sử dụng theo một hoặc nhiều đường sử dụng bằng cách sử dụng một hoặc nhiều phương pháp khác nhau đã biết trong lĩnh vực. Người có hiểu biết trung bình trong lĩnh vực hiểu rõ rằng, đường dùng và/hoặc phương thức dùng sẽ thay đổi tùy thuộc vào kết quả mong muốn. Các đường sử dụng đối với protein được thiết kế theo sáng chế bao gồm đường tiêm tĩnh mạch, tiêm bắp, trong da, trong phúc mạc, dưới da, tủy sống hoặc đường tiêm khác, ví dụ bằng cách tiêm hoặc truyền. Cụm từ ‘dùng ngoài đường tiêu hóa' như được sử dụng ở đây có nghĩa là các phương thức sử dụng không phải trong ruột và dùng tại chỗ, thường bằng đường tiêm, và bao gồm nhưng không bị giới hạn ở, tiêm tĩnh mạch, tiêm bắp, trong động mạch, trong khoang, trong bao, trong ổ mắt, trong tim, trong da, trong phúc mạc, qua khí quản, tiêm và truyền dưới da, dưới màng cứng, nội nhãn, dưới bao, dưới khoang nhện, trong tủy sống, ngoài màng cứng và trong mô phân sinh.

Theo cách khác, protein dung hợp trị liệu theo sáng chế có thể được sử dụng theo đường không phải ngoài đường tiêu hóa, như đường dùng tại chỗ, biểu bì hoặc niêm mạc.

Các protein dung hợp trị liệu theo sáng chế có thể được điều chế với chất mang sẽ bảo vệ các protein này chống lại sự giải phóng nhanh, như chế phẩm giải phóng có kiểm soát, bao gồm các mô cấy, miếng dán thẩm thấu qua da và hệ thống phân phối được bao vi nang. Polyme tương thích sinh học, có thể thoái biến sinh học có thể được sử dụng, như etylen vinyl axetat, polyanhydrit, axit polyglycolic, colagen, polyorthoeste, và axit polylactic. Các phương pháp điều chế các chế phẩm như vậy đã được cấp bằng sáng chế hoặc thường được người có hiểu biết trung bình trong lĩnh vực kỹ thuật này biết đến. Xem, ví dụ, Sustained and Controlled Release Drug Delivery Systems, J.R. Robinson, ed., Marcel Dekker, Inc., New York, 1978.

Theo một số phương án nhất định, các protein dung hợp trị liệu theo sáng chế có thể được phối chế để đảm bảo sự phân bố thích hợp *in vivo*. Ví dụ, hàng rào máu não (BBB) loại trừ nhiều hợp chất ưa nước cao. Để đảm bảo rằng hợp chất trị liệu theo sáng chế vượt qua BBB (nếu muốn), chúng có thể được tạo chế phẩm, ví dụ như, trong liposom. Đối với các phương pháp sản xuất liposom, xem, ví dụ, các bằng sáng chế Mỹ số 4,522,811; 5,374,548; và 5,399,331. Các liposom có thể bao gồm một hoặc nhiều gốc được vận chuyển có chọn lọc vào các tế bào hoặc cơ quan cụ thể, do đó tăng cường phân phối thuốc theo mục tiêu (xem, ví dụ, Ranade VV (1989) J. Clin. Pharmacol., 29: 685).

Sử dụng và phương pháp trị liệu theo sáng chế

Các protein dung hợp trị liệu theo sáng chế có các ứng dụng trị liệu và chẩn đoán *in vitro* và *in vivo*. Ví dụ, các phân tử này có thể được sử dụng cho các tế bào trong nuôi cấy, ví dụ, *in vitro*, hoặc ở đối tượng, ví dụ, *in vivo*, để điều trị, ngăn ngừa hoặc chẩn đoán nhiều loại rối loạn. Các phương pháp này đặc biệt thích hợp để điều trị, ngăn ngừa hoặc chẩn đoán viêm cấp tính hoặc mạn tính và các rối loạn vi mạch và cơ quan do hệ thống miễn dịch điều khiển.

Các protein dung hợp trị liệu theo sáng chế, trong khi không bị giới hạn, rất hữu ích cho việc điều trị, phòng ngừa hoặc cải thiện các tổn thương cơ quan viêm cấp tính và mãn tính, đặc biệt là các tổn thương viêm trong đó cơ chế thanh thải nội môi nội sinh hoặc các con đường sinh tế bào để loại bỏ các tế bào chết, các mảnh tế bào và vi hạt tiền

huyết khối/tiền viêm được điều hòa giảm đáng kể. Ví dụ về các tổn thương cơ quan viêm cấp tính bao gồm nhồi máu cơ tim, tổn thương thận cấp (AKI), đột quỵ và viêm cấp tính và chấn thương cơ quan do thiếu máu cục bộ/tái tưới máu như thiếu máu cục bộ/tái tưới máu đường tiêu hóa, gan, lá lách, phổi, thận, tuyến tụy, tim, não, tủy sống và/hoặc chi bị dập nát.

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện việc ức chế hoặc làm chậm quá trình đông máu, điều trị vi sinh vật, bệnh viêm ruột (IBD), hấp thu axit béo và/hoặc giảm nhu động dạ dày, rối loại phụ thuộc vi khuẩn, xơ vữa động mạch, sửa chữa tim, xơ hóa mô, tổn thương gan cấp, bệnh gan mãn tính, viêm gan nhiễm mỡ không do rượu (NASH), bệnh mạch máu, rối loạn mạch máu do tuổi tác, bệnh đường ruột, nhiễm trùng huyết, rối loạn xương, ung thư, tan máu bẩm sinh, viêm tụy, viêm gan, viêm nội tâm mạc, viêm phổi, tổn thương phổi cấp, viêm xương khớp, viêm nha chu, viêm do chấn thương mô, viêm đại tràng, tiểu đường, sốc xuất huyết, thải ghép, tổn thương do phóng xạ, lách to, AKI do nhiễm trùng máu hoặc suy đa cơ quan, bỏng cấp, hội chứng suy hô hấp ở người lớn và trẻ em, chữa lành vết thương, phục hồi gân và các bệnh thần kinh.

Trong một phương án, các bệnh thần kinh có thể được lựa chọn từ các tình trạng có thành phần bệnh lý thần kinh, viêm thần kinh và/hoặc thoái hóa thần kinh bao gồm các triệu chứng như hội chứng ốm đau, buồn nôn, né tránh thụ động, ức chế hành vi nhanh nhẹn, rối loạn trí nhớ và rối loạn chức năng ghi nhớ. Ví dụ về các bệnh thần kinh bao gồm các bệnh thần kinh liên quan đến amyloid-beta như bệnh Alzheimer, bệnh Parkinson và trầm cảm.

Theo một phương án, các rối loạn về xương có thể được lựa chọn từ các tình trạng bao gồm loãng xương, nhuyễn xương, xơ cứng và hoại tử xương. Cụ thể hơn là, việc sử dụng protein dung hợp theo sáng chế có thể ức chế sự biểu hiện của ít nhất một đánh dấu tế bào hủy xương, như NFATc1, cathepsin K và αvβ3 integrin. Trong một phương án, việc sử dụng này ức chế quá trình tạo tế bào xương. Trong một phương án khác, việc sử dụng này ức chế quá trình tạo xương do RANKL gây ra. Trong một phương án khác nữa, việc sử dụng này ức chế quá trình tiêu xương. Trong một phương án khác nữa, việc sử dụng này ức chế sự biểu hiện của ít nhất một chất kích thích tiêu xương, như chất kích thích tiêu xương bao gồm TNF, IL-6, IL-17A, MMP-9, Ptgs2, RANKL, Tnfsf11, CXCL1, CXCL2, CXCL3, CXCL5 và các kết hợp của chúng. Trong một

phương án khác, việc sử dụng này ức chế sự biểu hiện của ít nhất một xytokin gây viêm được chọn từ nhóm bao gồm IL-8 và CCL2/MCP-1.

Trong một phương án, xơ hóa mô có thể là xơ hóa ở gan, phổi, cơ hoành, thận, não, tim, trong đó protein dung hợp theo sáng chế làm giảm biểu hiện colagen. Theo một phương án, xơ hóa phổi là xơ hóa phổi mô kẽ (IPF). Theo một phương án, xơ hóa gan là xơ gan, có thể do NASH hoặc không.

Nhiều bệnh đường hô hấp có đặc điểm là tích tụ các tế bào chết theo chương trình. Hơn nữa, quá trình tiêu hủy tế bào chết và sự thực bào của đại thực bào trong rối loạn phổi tắc nghẽn mãn tính (COPD) có liên quan đến các đợt cấp và mức độ nghiêm trọng. Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện các bệnh đường hô hấp, như hội chứng suy hô hấp cấp hoặc COPD. Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện tổn thương phổi cấp (ALI), ví dụ, tổn thương phổi do hít hoặc hút các hợp chất hoặc thuốc ngoại sinh hoặc nội sinh độc hại; chấn thương phổi do phù phổi, sốc, viêm tụy, bỏng, chấn thương lồng ngực hoặc đa chấn thương, bức xạ, nhiễm trùng máu, mầm bệnh (vi khuẩn, virut hoặc ký sinh trùng như bệnh co thắt); các bệnh suy phổi mãn tính dẫn đến giảm oxy máu.

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của tổn thương phổi do virut thuộc loại Cornona gây ra, ví dụ, SARS-CoV, SARS-CoV-2 hoặc MERS-CoV. Trong một phương án, các protein dung hợp trị liệu theo sáng chế được cung cấp để sử dụng trong việc điều trị nhiễm SARS-CoV-2 ở các bệnh nhân COVID 19.

Các protein dung hợp điều trị theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của suy phổi liên quan đến truyền máu (TRALI).

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của các bệnh suy phổi mãn tính dẫn đến giảm oxy máu.

Các protein dung hợp trị liệu theo sáng chế, ví dụ, các protein dung hợp trị liệu chứa miền EDIL3 theo sáng chế, cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của sự dính phúc mạc sau phẫu thuật.

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của suy tim.

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của chạy thận nhân tạo.

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của chức năng mảnh ghép bị trì hoãn hoặc của bệnh mảnh ghép chống lại vật chủ.

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của bệnh tê cóng nghiêm trọng, bàn chân rãnh, viêm da mủ hạch/hoại thư.

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của bệnh lý do vi khuẩn, nấm, virut hoặc ký sinh trùng gây ra (ví dụ, nhiễm trùng máu hoặc các bệnh lý khác do mầm bệnh trực tiếp gây ra như bệnh than, bệnh dịch hạch, nhiễm trùng mô mềm hoại tử (NSTI như viêm cân hoại tử,) viêm tủy xương, sốt rét).

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của chấn thương/đa chấn thương do tai nạn gây thương tích, như tai nạn lao động, ngã, tai nạn giao thông, chấn thương đạn đạo và chiến đấu hoặc các cơ chế chấn thương khác.

Các protein dung hợp trị liệu theo sáng chế cũng có thể hữu ích cho việc chẩn đoán, điều trị, phòng ngừa hoặc cải thiện mức độ nghiêm trọng của bệnh lý trung gian hủy xương.

Các protein dung hợp trị liệu theo sáng chế có thể được sử dụng làm thành phần hoạt tính duy nhất hoặc kết hợp với, ví dụ, như chất bổ trợ cho hoặc kết hợp với, các loại thuốc khác, ví dụ, các chất ức chế miễn dịch hoặc điều hòa miễn dịch hoặc các chất chống viêm khác hoặc ví dụ, chất gây độc tế bào hoặc chống ung thư, ví dụ để điều trị hoặc phòng ngừa các bệnh nêu trên.

Được sử dụng 'kết hợp', liên quan đến chất trị liệu bổ sung, có nghĩa là hai (hoặc nhiều) phương pháp điều trị khác nhau được cung cấp cho đối tượng trong quá trình đối tượng mắc chứng rối loạn, ví dụ, hai hoặc nhiều phương pháp điều trị được phân phối

sau đối tượng đã được chẩn đoán mắc chứng rối loạn và trước khi chứng rối loạn được chữa khỏi hoặc loại bỏ hoặc việc điều trị đã chấm dứt vì những lý do khác. Theo một số phương án, sự phân phối của một việc điều trị vẫn xảy ra khi sự phân phối thứ hai bắt đầu, sao cho có sự gối lên của việc dùng. Điều này đôi khi được đề cập đến ở đây là "đồng thời" hoặc "sự phân phối cùng lúc". Theo phương án khác, sự phân phối của một việc điều trị kết thúc trước khi việc phân phối của việc điều trị khác bắt đầu. Theo một số phương án của một trong hai trường hợp, việc điều trị hữu hiệu hơn nhờ việc dùng kết hợp. Ví dụ, việc điều trị thứ hai hiệu quả hơn, ví dụ, hiệu quả tương đương được thấy với ít hơn việc điều trị thứ hai, hoặc việc điều trị thứ hai làm giảm các triệu chứng ở mức độ lớn hơn sẽ thấy nếu việc điều trị thứ hai được thực hiện mà không có việc điều trị đầu tiên, hoặc tình trạng tương tự xảy ra với việc điều trị đầu tiên. Theo một số phương án, việc phân phối nhờ vậy mang lại sự giảm đi của triệu chứng, hoặc thông số khác liên quan đến rối loạn lớn hơn so với thông số quan sát được với một việc điều trị được phân phối khi không có việc điều trị khác. Tác dụng của hai việc điều trị có thể cộng gộp một phần, cộng gộp toàn bộ, hoặc lớn hơn cộng gộp. Việc phân phối có thể là việc phân phối mà tác dụng của việc điều trị thứ nhất được phân phối vẫn có thể phát hiện được khi việc điều trị thứ hai được phân phối.

Thuật ngữ 'đồng thời' không giới hạn ở việc sử dụng các liệu pháp (ví dụ, thuốc dự phòng hoặc điều trị) cùng một lúc, mà nó có nghĩa là dược phẩm bao gồm các protein dung hợp trị liệu của chúng theo sáng chế được sử dụng cho đối tượng theo một trình tự và trong một khoảng thời gian sao cho các protein dung hợp có thể hoạt động cùng với (các) chất trị liệu bổ sung để mang lại lợi ích tăng lên so với khi chúng được sử dụng theo cách khác. Ví dụ, mỗi liệu pháp có thể được sử dụng cho đối tượng cùng lúc hoặc tuần tự theo bất kỳ thứ tự nào tại các thời điểm khác nhau; tuy nhiên, nếu không được sử dụng cùng lúc, chúng nên được dùng đủ gần về thời gian để mang lại hiệu quả điều trị hoặc dự phòng mong muốn. Mỗi liệu pháp có thể được dùng cho đối tượng một cách riêng rẽ, ở dạng thích hợp bất kỳ và theo đường thích hợp bất kỳ.

Protein dung hợp trị liệu như được mô tả ở đây và (các) chất trị liệu bổ sung có thể được sử dụng đồng thời, trong cùng một dược phẩm hoặc dược phẩm riêng biệt như protein dung hợp được bộc lộ, hoặc tuần tự. Đối với cách sử dụng tuần tự, protein dung hợp như được mô tả ở đây, có thể được sử dụng trước tiên và chất bổ sung có thể được sử dụng thứ hai, hoặc thứ tự sử dụng có thể được đảo ngược. (Các) chất trị liệu bổ sung

có thể được sử dụng cho đối tượng theo đường dùng giống hoặc khác so với protein dung hợp.

Protein dung hợp trị liệu như được mô tả ở đây và/hoặc (các) chất trị liệu bổ sung, các quy trình hoặc phương thức điều trị có thể được sử dụng trong giai đoạn rối loạn hoạt động, hoặc trong giai đoạn bệnh thuyên giảm hoặc ít hoạt động hơn. Protein dung hợp điều trị như được mô tả ở đây, có thể được sử dụng trước phương pháp điều trị khác, đồng thời với phương pháp điều trị, sau điều trị hoặc trong thời gian thuyên giảm rối loạn.

Khi được sử dụng kết hợp, protein dung hợp trị liệu như được mô tả ở đây và chất trị liệu bổ sung (ví dụ, chất thứ hai hoặc thứ ba), hoặc tất cả, có thể được sử dụng với lượng hoặc liều cao hơn, thấp hơn hoặc bằng với lượng hoặc liều dùng của từng chất được sử dụng riêng lẻ, ví dụ, như đơn trị liệu. Trong các phương án nhất định, protein dung hợp trị liệu như được mô tả ở đây, chất bổ sung (ví dụ, chất thứ hai hoặc thứ ba), hoặc tất cả, thấp hơn (ví dụ, ít nhất 20%, ít nhất 30%, ít nhất 40%, hoặc ít nhất 50%) so với số lượng hoặc liều dùng của từng chất được sử dụng riêng lẻ, ví dụ như trị liệu đơn. Theo các phương án khác, lượng hoặc liều dùng của protein dung hợp trị liệu như được mô tả ở đây, chất bổ sung (ví dụ, chất thứ hai hoặc thứ ba), hoặc tất cả, dẫn đến hiệu quả mong muốn (ví dụ, điều trị bệnh hoặc tình trạng viêm) là thấp hơn (ví dụ, ít nhất 20%, ít nhất 30%, ít nhất 40%, hoặc thấp hơn ít nhất 50%) so với lượng hoặc liều lượng của từng chất được sử dụng riêng lẻ, ví dụ như đơn trị liệu, cần thiết để đạt được cùng một hiệu quả điều trị.

Ví dụ, các protein dung hợp trị liệu theo sáng chế có thể được sử dụng kết hợp với DMARD, ví dụ, muối vàng, sulphasalazin, chống sốt rét, methotrexat, D-penicillamin, azathioprin, axit mycophenolic, tacrolimus, sirolimus, minocyclin, leflunomid, glucocorticoid; chất ức chế calcineurin, ví dụ xyclosporin A hoặc FK 506; bộ điều biến tuần hoàn tế bào lympho, ví dụ: FTY720 và FTY720 tương tự; chất ức chế mTOR, ví dụ: rapamycin, 40-O-(2-hydroxyethyl)-rapamycin, CCI779, ABT578, AP23573 hoặc TAFA-93; ascomycin có đặc tính ức chế miễn dịch, ví dụ ABT-281, ASM981, v.v.; thuốc corticosteroit; xyclophosphamit; azathioprin; leflunomit; mizoribin; mycophenolat mofetil; 15-deoxyspergualin hoặc chất tương đồng ức chế miễn dịch, chất tương tự hoặc dẫn xuất của chúng; kháng thể đơn dòng ức chế miễn dịch, ví dụ, kháng thể đơn dòng đối với các thụ thể bạch cầu, ví dụ, MHC, CD2, CD3,

CD4, CD7, CD8, CD25, CD28, CD40. CD45, CD58, CD80, CD86 hoặc các phối tử của chúng; các hợp chất điều hòa miễn dịch khác, ví dụ phân tử liên kết tái tổ hợp có ít nhất một phần miền ngoại bào của CTLA4 hoặc một thể đột biến của chúng, ví dụ ít nhất một phần ngoại bào của CTLA4 hoặc một phần đột biến của chúng liên kết với trình tự protein không phải CTLA4, ví dụ: CTLA4Ig (ví dụ: được chỉ định ATCC 68629) hoặc đột biến của chúng, ví dụ: LEA29Y; chất ức chế phân tử bám dính, ví dụ chất đối kháng LFA-1, chất đối kháng ICAM-1 hoặc -3, chất đối kháng VCAM-4 hoặc chất đối kháng VLA-4; hoặc chất hóa trị liệu, ví dụ: paclitaxel, gemcitabin, cisplatinum, doxorubicin hoặc 5-flouracil; chất chống TNF, ví dụ: kháng thể đơn dòng đối với TNF, ví dụ infliximab, adalimumab, CDP870, hoặc cấu trúc thụ thể thành TNF-RI hoặc TNF-RII, ví dụ: Etanercept, PEG-TNF-RI; chất phong bế xytokin tiền viêm, chất phong bế IL-1, ví dụ: Anakinra hoặc bẫy IL-1, canakinumab, chất phong bế IL-13, chất phong bế IL-4, chất phong bế IL-6; chất phong bế chemokin, ví dụ chất ức chế hoặc chất kích hoạt protcaza, ví dụ: metalloproteaza, kháng thể kháng IL-15, kháng thể kháng IL-6, kháng thể kháng IL-4, kháng thể kháng IL-13, kháng thể kháng CD20, NSAID, như aspirin hoặc chất chống nhiễm trùng; Các chất đối kháng kiểu phân tử liên quan đến tổn thương (DAMP) hoặc kiểu phân tử liên quan đến mầm bệnh (PAMP), ví dụ chất chuyển hóa, chất khử độc, chất loại bỏ, ví dụ: chất chuyển hóa ATP, chất điều biến HMGB-1, chất giải độc histon; chất ức chế các phản ứng miễn dịch gây ra siêu kháng nguyên; chất ức chế bổ thể và thiết bị di chuyển tế bào chất ngoài cơ thể.

Kit

Cũng trong phạm vi của sáng chế là kit chứa các chế phẩm, ví dụ, các protein dung hợp trị liệu theo sáng chế và hướng dẫn sử dụng. Kit này bao gồm lượng hữu hiệu trị liệu của protein dung hợp theo sáng chế. Ngoài ra, kit này có thể bao gồm các phương tiện để sử dụng protein dung hợp trị liệu (ví dụ, kim tiêm tự động, ống tiêm và lọ, ống tiêm đã được nạp sẵn, bút đã được nạp sẵn) và hướng dẫn sử dụng. Kit này có thể chứa các chất trị liệu bổ sung (được mô tả *sau đây*) để điều trị bệnh nhân mắc bệnh tự miễn dịch hoặc rối loạn viêm nhiễm hoặc AOI. Kit như vậy cũng có thể bao gồm hướng dẫn sử dụng của protein dung hợp trị liệu để điều trị bệnh nhân. Các hướng dẫn như vậy có thể cung cấp liều lượng, đường dùng, phác đồ và tổng thời gian điều trị khi sử dụng protein dung hợp kèm theo. Kit thường bao gồm nhãn cho biết mục đích sử dụng của các thành phần của kit. Thuật ngữ nhãn bao gồm bất kỳ văn bản hoặc tài liệu ghi lại nào

được cung cấp trên hoặc cùng với kit, hoặc những thứ khác đi kèm với kit. Kit còn có thể bao gồm các công cụ để chẩn đoán xem bệnh nhân có thuộc nhóm sẽ đáp ứng với điều trị bằng protein dung hợp trị liệu theo sáng chế hay không, như đã định nghĩa ở trên.

Các phương án thực hiện sáng chế

Sáng chế đề xuất các phương án sau đây:

1. Protein dung hợp trị liệu để tăng cường quá trình tiêu hủy tế bào chết bao gồm miền liên kết integrin, miền liên kết phosphatidylserin (PS) và miền hòa tan.

2. Protein dung hợp theo phương án 1, trong đó miền hòa tan:

(i) được liên kết với miền liên kết integrin;

(ii) được liên kết với miền liên kết PS;

(iii) được chèn ở giữa miền liên kết integrin và miền liên kết PS;

(iv) được chèn trong miền liên kết integrin; hoặc

(v) được chèn trong miền liên kết PS.

3. Protein dung hợp theo phương án 1 hoặc phương án 2, trong đó miền liên kết integrin liên kết với một hoặc nhiều integrin.

4. Protein dung hợp theo phương án 3, trong đó miền liên kết integrin liên kết với integrin αvβ3 và/hoặc αvβ5 và/hoặc α8β1.

5. Protein dung hợp theo phương án 3 hoặc phương án 4, trong đó miền liên kết integrin bao gồm môtip Arginin-Glyxin-axit Aspartic (RGD).

6. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó miền hòa tan được liên kết trực tiếp với miền liên kết integrin, với miền liên kết PS hoặc với cả hai miền.

7. Protein dung hợp theo phương án bất kỳ trong số các phương án từ 1 đến 6, trong đó miền hòa tan được liên kết trực tiếp với miền liên kết integrin và/hoặc miền liên kết PS bằng cầu nối.

8. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó miền hòa tan bao gồm albumin huyết thanh người (HSA), miền 3 của HSA (HSA D3), Fc-IgG, hoặc biến thể chức năng của nó.

9. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó miền hòa tan bao gồm albumin huyết thanh người (HSA), hoặc biến thể chức năng của nó.

7. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó miền liên kết integrin có trình tự axit amin của SEQ ID NO: 2, hoặc tương đồng trình tự ít nhất 90% với nó.

8. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó miền liên kết PS có trình tự axit amin của SEQ ID NO: 3, hoặc tương đồng trình tự ít nhất 90% với nó; hoặc miền liên kết PS có trình tự axit amin của SEQ ID NO: 76, hoặc tương đồng trình tự ít nhất 90% với nó.

9. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó miền hòa tan là HSA và có trình tự axit amin của SEQ ID NO: 4, hoặc tương đồng trình tự ít nhất 90% với nó.

10. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó miền liên kết integrin có trình tự axit amin của SEQ ID NO: 2, hoặc tương đồng trình tự ít nhất 90% với nó và miền liên kết PS có trình tự axit amin của SEQ ID NO: 78, hoặc tương đồng trình tự ít nhất 90% với nó.

11. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó miền liên kết integrin có trình tự axit amin của SEQ ID NO: 77, hoặc tương đồng trình tự ít nhất 90% với nó và miền liên kết PS có trình tự axit amin của SEQ ID NO: 3, hoặc của SEQ ID NO: 76, hoặc tương đồng trình tự ít nhất 90% với nó.

12. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó protein dung hợp này:

a. thúc đẩy quá trình tiêu hủy tế bào chết bởi tế bào nội mô ở thử nghiệm tiêu hủy tế bào chết của tế bào nội mô-tế bào Jurkat người;

b. khôi phục quá trình tiêu hủy tế bào chết bị suy giảm của các đại thực bào ở thử nghiệm tiêu hủy tế bào chết của bạch cầu trung tính bởi đại thực bào của người;

c. làm giảm số lượng vi hạt huyết tương bằng sự thanh thải trong thử nghiệm tiêu hủy tế bào chết của nội mô-vi hạt của người; và/hoặc

d. bảo vệ chống lại các thương tổn đa cơ quan ở mô hình tổn thương thận cấp

13. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, bao gồm trong trình tự: miền liên kết integrin-HSA-miền liên kết PS.

14. Protein dung hợp trị liệu bao gồm MFG-E8 và miền hòa tan, trong đó MFG-E8 bao gồm từ đầu cùng N đến đầu cùng C: miền giống EGF, miền C1 và miền C2, và bao gồm trình tự từ MFG-E8 người kiểu dại (SEQ ID NO: 1), hoặc MFG-E8 có SEQ ID NO: 75, hoặc biến thể chức năng của nó.

15. Protein dung hợp theo phương án 14, trong đó miền hòa tan được chèn ở giữa miền giống EGF và miền C1.

16. Protein dung hợp theo phương án 14 hoặc phương án 15, trong đó miền hòa tan là HSA, HSA D3 hoặc Fc-IgG, hoặc biến thể chức năng của nó.

17. Protein dung hợp theo phương án bất kỳ trong số các phương án từ 1 đến 16, trong đó protein được thiết kế có trình tự axit amin của SEQ ID NO: 42, hoặc tương đồng trình tự ít nhất 90% với nó.

18. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó protein dung hợp này có trình tự axit amin của SEQ ID NO: 44, hoặc tương đồng trình tự ít nhất 90% với nó; hoặc SEQ ID NO: 47, hoặc tương đồng trình tự ít nhất 90% với nó; hoặc SEQ ID NO: 48, hoặc tương đồng trình tự ít nhất 90% với nó.

19. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó protein dung hợp này có trình tự axit amin của SEQ ID NO: 80, hoặc tương đồng trình tự ít nhất 90% với nó.

20. Protein dung hợp theo phương án bất kỳ trong số các phương án nêu trên, trong đó protein dung hợp này có trình tự axit amin của SEQ ID NO: 82, hoặc tương đồng trình tự ít nhất 90% với nó.

21. Axit nucleic được phân lập mã hóa trình tự axit amin theo phương án bất kỳ trong số các phương án từ 17 đến 20.

22. Vectơ tách dòng hoặc biểu hiện chứa axit nucleic theo phương án 21.

23. Vectơ virut chứa axit nucleic được phân lập theo phương án 21, tốt hơn là vectơ virut này chứa axit nucleic được phân lập theo phương án 21 thu được từ AAV.

24. Vectơ virut theo phương án 23, trong đó vectơ này được sử dụng cho đối tượng, ví dụ, đối tượng người, cần sử dụng.

25. Vectơ virut theo phương án 23, để sử dụng trong điều trị và/hoặc ngăn ngừa bệnh như được liệt kê ở đây.

26. Tế bào chủ tái tổ hợp để sản xuất protein dung hợp trị liệu, bao gồm một hoặc nhiều vectơ tách dòng hoặc biểu hiện theo phương án 22 và tùy ý, các tín hiệu tiết.

27. Tế bào chủ tái tổ hợp theo phương án 26, trong đó tế bào chủ này là, ví dụ tế bào nhân sơ, tế bào nấm men, tế bào côn trùng hoặc tế bào động vật có vú.

28. Protein dung hợp theo phương án bất kỳ trong số các phương án từ 1 đến 20, trong đó sự biểu hiện của protein ở tế bào chủ tạo ra năng suất là ít nhất 10 mg/L.

29. Protein dung hợp theo phương án bất kỳ trong số các phương án từ 1 đến 20, trong đó sự biểu hiện của protein ở tế bào động vật có vú tạo ra sự tăng năng suất ít nhất 100 lần so với MFG-E8 kiểu dại (SEQ ID NO: 1).

30. Dược phẩm chứa protein dung hợp theo phương án bất kỳ trong số các phương án từ 1 đến 20 và ít nhất một chất mang dược dụng.

31. Phương pháp điều trị hoặc ngăn ngừa rối loạn viêm hoặc tổn thương cơ quan do viêm ở đối tượng cần điều trị, bao gồm bước dùng cho đối tượng này lượng có hiệu quả điều trị của protein dung hợp theo phương án bất kỳ trong số các phương án từ 1 đến 20.

32. Protein dung hợp theo phương án bất kỳ trong số các phương án từ 1 đến 20 để sử dụng trong điều trị hoặc ngăn ngừa rối loạn viêm hoặc tổn thương cơ quan do viêm ở đối tượng cần điều trị.

33. Phương pháp theo phương án 31 hoặc sử dụng theo phương án 32, trong đó rối loạn viêm hoặc tổn thương cơ quan do viêm là tổn thương thận cấp, nhiễm trùng máu, nhồi máu cơ tim, đột quỵ cấp, bỏng, tổn thương do chấn thương, và viêm và tổn thương cơ quan do thiếu máu cục bộ/tái tưới máu.

34. Phương pháp theo phương án 31 hoặc sử dụng theo phương án 32, trong đó rối loạn viêm hoặc tổn thương cơ quan do viêm là tổn thương thận cấp.

35. Phương pháp theo phương án 31 hoặc sử dụng theo phương án 32, trong đó rối loạn viêm hoặc tổn thương cơ quan do viêm là nhồi máu cơ tim,

36. Phương pháp theo phương án 31 hoặc sử dụng theo phương án 32, trong đó rối loạn viêm hoặc tổn thương cơ quan do viêm là đột quỵ,

37. Phương pháp theo phương án 31 hoặc sử dụng theo phương án 32, trong đó rối loạn viêm hoặc tổn thương cơ quan do viêm là tổn thương phổi cấp (ví dụ hội chứng suy hô hấp cấp) hoặc tổn thương gan hoặc tổn thương ruột cấp.

38. Phương pháp theo phương án 31 hoặc sử dụng theo phương án 32, trong đó protein dung hợp được sử dụng kết hợp với chất trị liệu khác.

39. Phương pháp hoặc sử dụng theo phương án 38, trong đó chất trị liệu khác này là chất ức chế miễn dịch, chất điều biến miễn dịch, chất chống viêm, chất chống oxy hóa, chất chống nhiễm trùng, chất gây độc tế bào hoặc chất chống ung thư.

40. Protein dung hợp trị liệu bao gồm MFG-E8 và miền hòa tan, trong đó MFG-E8 bao gồm từ đầu cùng N đến đầu cùng C: miền giống EGF, miền C1 và miền C2, và bao gồm trình tự từ MFG-E8 người kiểu dại (SEQ ID NO: 1), hoặc của SEQ ID NO: 75, hoặc biến thể chức năng của nó.

41. Protein dung hợp theo phương án 40, trong đó miền hòa tan được liên kết với đầu cùng N hoặc đầu cùng C của MFG-E8 (SEQ ID NO: 1 hoặc SEQ ID NO: 75).

42. Protein dung hợp theo phương án 40, trong đó miền hòa tan được chèn ở giữa miền giống EGF và miền C1.

43. Protein dung hợp theo phương án 41, trong đó miền hòa tan được chèn ở giữa miền C1 và miền C2.

44. Protein dung hợp theo phương án bất kỳ trong số các phương án từ 40 đến 43, trong đó miền hòa tan là HSA, HSA D3 hoặc Fc-IgG, hoặc biến thể chức năng của nó.

38. Axit nucleic được phân lập mã hóa protein dung hợp theo phương án bất kỳ trong số các phương án từ 33 đến 37.

39. Vectơ tách dòng hoặc biểu hiện chứa axit nucleic theo phương án 38.

40. Vectơ virut chứa axit nucleic được phân lập theo phương án 38, tốt hơn là vectơ virut này chứa axit nucleic được phân lập theo phương án 38 thu được từ AAV.

41. Vectơ virut theo phương án 40, trong đó vectơ này được sử dụng cho đối tượng, ví dụ, đối tượng người, cần sử dụng.

42. Vectơ virut theo phương án 40, để sử dụng trong điều trị và/hoặc ngăn ngừa bệnh như được liệt kê ở đây.

43. Tế bào chủ tái tổ hợp để sản xuất protein dung hợp trị liệu, bao gồm một hoặc nhiều vectơ tách dòng hoặc biểu hiện theo phương án 39 và tùy ý, các tín hiệu tiết.

44. Tế bào chủ tái tổ hợp theo phương án 43, trong đó tế bào chủ này là, ví dụ tế bào nhân sơ, tế bào nấm men, tế bào côn trùng hoặc tế bào động vật có vú.

45. Protein dung hợp theo phương án bất kỳ trong số các phương án từ 33 đến 37, trong đó sự biểu hiện của protein ở tế bào chủ tạo ra năng suất là ít nhất 10 mg/L.

46. Protein dung hợp theo phương án bất kỳ trong số các phương án từ 33 đến 37, trong đó sự biểu hiện của protein trong tế bào động vật có vú dẫn đến sự tăng năng suất ít nhất 100 lần so với MFG-E8 kiểu dại.

Cần hiểu là mỗi phương án có thể được kết hợp với một hoặc nhiều phương án khác, đến mức độ mà hỗn hợp này tạo ra hợp chất ổn định và phù hợp với phần mô tả các phương án này. Cần hiểu thêm là các phương án được đề xuất trên đây được hiểu là bao gồm tất cả các phương án, bao gồm các phương án thu được từ các tổ hợp của các phương án.

Tất cả các tài liệu tham khảo được trích dẫn ở đây, bao gồm sáng chế, đơn sáng chế, giấy tờ, công bố, sách văn bản, và tương tự, và các tài liệu tham khảo được trích dẫn ở đây, trong phạm vi chưa có, đều được kết hợp ở đây bằng cách viện dẫn toàn bộ.

Ví dụ thực hiện sáng chế

Các ví dụ sau đây được cung cấp để minh họa thêm cho sáng chế nhưng không giới hạn phạm vi của nó. Các biến thể khác của sáng chế sẽ dễ thấy đối với người có trình độ trung bình trong lĩnh vực và được bao hàm bởi các điểm yêu cầu bảo hộ kèm theo.

Ví dụ 1: Tạo ra các protein dung hợp

MFG-E8 là protein đa miền gồm miền yếu tố sinh trưởng biểu bì (giống EGF) đầu cùng N và hai miền C kiểu lectin đầu cùng C (C1 và C2). Các nỗ lực tạo ra protein người có chiều dài đầy đủ tái tổ hợp, như được ghi lại trong tài liệu, đã chỉ ra rằng các sự kết tụ và tỷ lệ biểu hiện protein là rất thấp (Castellanos *et al.,* (2016) Protein Expression Purification 1124: 10-22). Do đó, để cố gắng hòa tan protein và tăng cường sự biểu hiện của nó, các tác giả sáng chế đã nghiên cứu ảnh hưởng của việc dung hợp một số protein vào MFG-E8.

Miền hòa tan (SD) có nguồn gốc từ human Fc-IgG1, albumin huyết thanh người (HSA) và miền 3 của HSA (HSA D3) được dung hợp ở các vị trí khác nhau với MFG-E8; tại đầu cùng N hoặc C, hoặc ở giữa các miền EGF và C1 hoặc C1 và C2, như được thể hiện bằng sơ đồ trên Hình 1. Hơn nữa, sự dung hợp với Fc-IgG1 hoặc HSA có khả năng kéo dài thời gian bán hủy của phân tử *in vivo*, vì những protein này liên kết với FcRn. Sự dung hợp của MFG-E8 với Fc-IgG1 hoặc HSA cũng có thể tăng cường việc sản xuất và độ hòa tan (Castellanos *et al.*, (2016) *nêu trên*) của protein dung hợp như được thể hiện trong các ví dụ sau đây.

Bảng 5 thể hiện việc liên kết của protein dung hợp FP330 (EGF-HSA-C1-C2; SEQ ID NO: 42) bao gồm đoạn chèn HSA, với thụ thể Fc sơ sinh của người (Cũng xem Ví dụ 5.1).

Bảng 5: Ái lực liên kết của protein dung hợp FP330 với FcRn người

Mẫu	FcRn người (pH 5,8)		FcRn người (pH 7,4)	
	K_D (nM)	n	K_D (nM)	n
FP330 (SEQ ID NO: 42)	1380 ± 95	2	không liên kết	1

Ví dụ 2: Tạo ra thể dung hợp wtMFG-E8 và MFG-E8 HSA; sự biểu hiện và tinh chế

Các phương pháp tạo ra protein dung hợp được mô tả dưới đây; một cách ngắn gọn, thể dung hợp MFG-E8 và MFG-E8 và thể dung hợp EDIL, đặc biệt là sự dung hợp với HSA, được tạo ra theo phương pháp sau đây.

ADN được tổng hợp tại GeneArt (Regensburg, Đức) và được tách dòng vào vectơ biểu hiện động vật có vú bằng cách sử dụng kỹ thuật tách dòng dựa trên sự gắn kết bằng enzym cắt giới hạn. Plasmit thu được được chuyển nạp vào trong các tế bào HEK293T. Đối với sự biểu hiện tạm thời của protein, các vectơ cho các chuỗi kiểu dại hoặc được thiết kế được chuyển nạp vào các tế bào HEK293T được làm thích nghi với huyền phù bằng cách sử dụng Polyetylenimin (PEI; Cat# 24765 Polysciences, Inc.). Thông thường, 100 ml tế bào ở dạng huyền phù với mật độ 1-2 triệu tế bào mỗi ml được chuyển nạp bằng ADN chứa 100 µg vectơ biểu hiện mã hóa các chuỗi được thiết kế này. Sau đó, các vectơ biểu hiện tái tổ hợp được đưa vào các tế bào chủ và cấu trúc được tạo ra bằng cách nuôi cấy thêm các tế bào trong thời gian 7 ngày để cho phép tiết vào môi trường nuôi cấy (HEK, môi trường không chứa huyết thanh) bổ sung axit pluronic 0,1%, glutamin 4mM và 0,25 µg/ml kháng sinh.

Các cấu trúc được tạo ra sau đó được tinh chế từ phần nổi không chứa tế bào, sử dụng sắc ký ái lực ion kim loại được cố định (IMAC), hoặc sự bắt giữ Protein A, hoặc sắc ký bắt giữ kháng HSA.

Khi protein được gắn thẻ his được IMAC bắt giữ, môi trường có điều kiện được lọc được trộn với nhựa IMAC (GE Healthcare), được cân bằng với 1% triton và NaPO4 20mM, NaCl 0,5Mn, Imidazol 20mM, độ pH 7,0. Nhựa được rửa ba lần với 15 thể tích cột gồm NaPO4 20mM, NaCl 0,5Mn, Imidazol 20mM, độ pH 7,0 trước khi protein được rửa giải bằng 10 thể tích cột đệm rửa giải (NaPO4 20mM, NaCl 0,5Mn, Imidazol 20mM, độ pH 7,0).

Khi protein được bắt giữ bằng sắc ký Protein A hoặc kháng HSA, môi trường có điều kiện được lọc được trộn với nhựa Protein A (CaptivA PriMab™, Repligen) hoặc nhựa kháng HSA (nền ái lực Albumin người chọn lọc bắt giữ, Thermo), được cân bằng với PBS, độ pH 7,4. Nhựa được rửa ba lần với 15 thể tích cột PBS, độ pH 7,4 trước khi protein được rửa giải với 10 thể tích cột đệm rửa giải (xitrat 50mM, NaCl 90mM, độ pH 2,5) và độ pH được trung hòa bằng cách sử dụng TRIS độ pH 10,0 1M.

Cuối cùng, các phân đoạn rửa giải được đánh bóng bằng cách sử dụng sắc ký loại cỡ (HiPrep Superdex 200, 16/60, GE Healthcare Life Sciences) và được phân tích bằng SDS-PAGE dựa trên chỉ thị chuẩn không nhuộm protein Precision plus (Biorad, ref # 161-0363).

Các gel biểu hiện đại diện cho các protein dung hợp được thể hiện trên Hình 2: Hình 2A: protein EGF-HSA-C1-C2 (FP330; SEQ ID NO: 42); Hình 2B: EGF-HSA-C1-C2 của protein EDIL3 (FP050; SEQ ID NO: 12); Hình 2C: protein không được khử và được khử EGF-Fc(KiH) C1-C2. Protein này là heterodime của FP071 (EGF-Fc(u)-C1-C2; SEQ ID NO: 18) với lỗ Fc-IgG1 (SEQ ID NO: 10); Hình 2D: protein EGF-HSA-C1 (FP260; SEQ ID NO: 34). Protein trong các điều kiện được khử và không được khử được thể hiện trên Hình 2C vì các heterodime có xu hướng phân rã trong các điều kiện khử do đó cả hai điều kiện đều được thử nghiệm. Các kết quả của sự biểu hiện và năng suất sau khi tinh chế đối với tập hợp các protein dung hợp khác được thể hiện trong Bảng 6; Như có thể thấy từ dữ liệu biểu hiện, các thể dung hợp HSA của MFG-E8, ngay cả với HSA ở các vị trí khác nhau, cho thấy sự cải thiện ít nhất 100 lần sự biểu hiện so với wtMFG-E8. Như được thể hiện trong cột bên phải của Bảng 6, thể dung hợp HSA của MFG-E8 cũng cho thấy năng suất tăng ít nhất 100 lần so với wtMFG-E8.

Bảng 6: Sự biểu hiện và năng suất của các protein dung hợp được biểu hiện trong dòng tế bào HEK

Protein	**Sự biểu hiện sau khi bắt His (mg/l)**	**Năng suất cuối cùng sau His và SEC (mg/l)**
wtMFG-E8	0,2	0,04
FP220 (HSA-EGF-C1-C2)	23	5,5
FP110 (EGF-C1-C2-HSA)	34	7,8
FP330 (EGF-HSA-C1-C2)	23	4,0

Các ví dụ khác của các protein dung hợp trị liệu theo sáng chế được tạo ra theo phương pháp trên và được phân tích thêm bằng SDS-PAGE (điện di trên gel natri dodexyl sulfat polyacrylamit), là các protein được tách ra dựa trên khối lượng phân tử của chúng. Mỗi protein được trộn với đệm Laemmli trước khi tải trên gel polyacrylamit (Biorad, 4-20% Mini-PROTEAN TGX Stain free). Sau 30 phút di chuyển ở 200V trong bộ đệm chạy TRIS-Glycin-SDS, các protein có trong gel đã được hiện lên trong máy ảnh kích hoạt không nhuộm (Biorad, Gel Doc EZ). Như được mô tả trên Hình 2E, SDS-PAGE cho thấy các protein tái tổ hợp đã được sản xuất và được tinh chế:

Dòng 1, 12:	Chỉ thị khối lượng phân tử (Biorad, protein Precision plus)		
Dòng 2:	His6_EGF[MFG-E8]_C1[MFG-E8]		23,87kDa
Dòng 3:	EGF[MFG-E8]_C1[MFG-E8]_His6	SEQ ID 115	23,87kDa
Dòng 4:	EGF[MFG-E8]_HSA_C1[MFG-E8]	SEQ ID 117	90,38kDa
Dòng 5:	EGF[MFG-E8]_HSA_C1[MFG-E8]	SEQ ID 74	89,27kDa
Dòng 6:	EGF[MFG-E8]_HSA_C1[MFG-E8]	SEQ ID 73	88,72kDa
Dòng 7:	EGF[EDIL3]_HSA_C1[EDIL3]	SEQ ID 71	98,22kDa
Dòng 8:	EGF[EDIL3]_HSA_C2[EDIL3]	SEQ ID 135	98,20kDa
Dòng 9:	EGF[MFG-E8]_HSA_C2[MFG-E8]	SEQ ID 137	88,45kDa
Dòng 10:	EGF[EDIL3]_HSA_C1_C2[MFG-E8]	SEQ ID 80	115,67kDa
Dòng 11:	EGF[MFG-E8]_HSA_C1_C2[EDIL3]	SEQ ID 82	107,32kDa

Ví dụ 3: Xác định đặc điểm của các protein được thiết kế MFG-E8-HSA

3.1 *Liên kết phosphatidylserin (hóa sinh)*

L-α-phosphatidylserin (não, lợn, Avanti 840032, Alabama, Mỹ) được hòa tan trong cloroform, được pha loãng trong metanol và được phủ lên các đĩa vi phiếm 384 giếng (Corning™ 3653, Kennebunk ME, Mỹ) ở 1 µg/mL. Sau khi ủ qua đêm ở 4°C, dung môi được làm bay hơi bằng cách sử dụng hệ thống SpeedVac™ (Thermo Scientific™). Các đĩa được xử lý bằng nước muối đệm phosphat (PBS) chứa 3% albumin huyết thanh bò (BSA) không có axit béo ở nhiệt độ phòng trong 1,5 giờ.

Việc liên kết của các protein dung hợp với L-α-phosphatidylserin được đánh giá bằng cách cạnh tranh với sự liên kết của MFG-E8/lactadherin chuột được biotin hóa (được sản xuất nội bộ, mMFG-E8:biotin). Các protein được pha loãng trong PBS chứa 3% BSA không có axit béo, độ pH 7,4 và được ủ với các đĩa vi phiếm được phủ L-α-phosphatidylserin trong 30 phút. mMFG-E8:biotin trong PBS chứa 3% BSA không chứa axit béo, độ pH 7,4 được bổ sung vào ở 1 nM và ủ thêm 30 phút. MMFG-E8:biotin không liên kết được loại bỏ qua ba bước rửa với dung dịch đệm rửa thử nghiệm miễn dịch huỳnh quang lanthanit được tăng cường phân ly (DELFIA™) (Perkin Elmer 1244-114 MA, Mỹ). Streptavidin được đánh dấu bằng Europium (Perkin Elmer 1244-360, Wallac Oy, Phần Lan) được bổ sung vào dung dịch đệm thử nghiệm DELFIA™ (Perkin Elmer 1244-111 MA, Mỹ) ở nhiệt độ phòng trong 20 phút. Tiếp theo là ba bước rửa với

dung dịch đệm thử nghiệm DELFIA™. Europium được đưa ra theo hướng dẫn của nhà sản xuất (Perkin Elmer 1244-105, Boston MA, Mỹ). Độ huỳnh quang phân giải theo thời gian của Europium được định lượng bằng thiết bị đọc đãi đa nhãn Envision™2103, Perkin Elmer, CT, Mỹ). Phân tích dữ liệu được thực hiện bằng phần mềm MS Excel và GraphPad Prism.

Các đĩa polypropylen là các đĩa vi phiếm liên kết protein thấp thường được sử dụng trong các phòng thí nghiệm để pha loãng nối tiếp. So với polystyren, các đĩa này có ưu điểm là giảm sự hao hụt protein trong quá trình pha loãng và thường được phân loại là các đĩa "liên kết protein thấp". Khi các dung dịch pha loãng của wtMFG-E8 được tạo ra trong các đĩa polypropylen, so với các dung dịch pha loãng được tạo ra trong các đĩa không liên kết, wtMFG-E8 bị mất hiệu lực trong thử nghiệm cạnh tranh L-α-phosphatidylserin. Những dữ liệu này, như được thể hiện trên Hình 3, gợi ý rằng wtMFG-E8 bị mất một phần trong các bước xử lý và pha loãng chất lỏng khi sử dụng các đĩa polypropylen đã được tối ưu hóa đối với sự liên kết protein thấp (Hình 3A). Những kết quả này chỉ ra rằng độ dính cố hữu của wtMFG-E8 đặt ra thách thức trong việc xử lý trong phòng thí nghiệm và rất có thể là trong quá trình tạo ra và sản xuất thuốc, nơi yêu cầu các bước bắt giữ và đánh bóng cần để tạo ra dược chất với năng suất cao và độ tinh khiết rất cao. Ngược lại, độ dính của protein được thiết kế FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) giảm mạnh so với wtMFG-E8 và hầu như không quan sát thấy có sự khác biệt nào giữa các độ pha loãng được thực hiện trong các đĩa không liên kết so với các đĩa polypropylen (Hình 3B). Những dữ liệu này gợi ý rằng việc chèn miền hòa tan vào các protein theo sáng chế có thể cải thiện việc xử lý kỹ thuật của chúng để cải thiện năng suất bước và do đó cải thiện năng suất tổng thể trong quá trình sản xuất.

Đánh giá sự liên kết của các protein dung hợp với L-α-phosphatidylserin được thể hiện trên Hình 4. Protein FP278 có nguồn gốc từ MFG-E8 được thiết kế (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) liên kết với PS được cố định và ở mức độ thấp hơn với phospholipit cardiolipin theo cách phụ thuộc nồng độ (Hình 4A). Sự liên kết của FP278 với L-α-phosphatidylserin được cố định hoặc liên kết với cardiolipin (1,3-bis (sn-3'-phosphatidyl)-sn-glyxerol) được phát hiện bằng cách sử dụng kháng thể chống lại miền EGF-L của wtMFG-E8. Cường độ liên kết của một số protein dung hợp tái tổ hợp với L-α-phosphatidylserin cố định được thể hiện trên Hình 4B. WtMFG-E8 người và các

protein dung hợp FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) và FP260 (EGF-HSA-C1; SEQ ID NO: 34) cạnh tranh hiệu quả với sự liên kết của 1 nM MFG-E8 chuột được biotin hóa với L-α-phosphatidylserin cố định theo cách phụ thuộc vào nồng độ. Các giá trị IC_{50} thu được đối với các protein dung hợp biểu thị các cường độ liên kết L-α-phosphatidylserin có độ tương tự cao của các miền C1-C2 của protein được thiết kế di truyền FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) so với wtMFG-E8 người. Bất ngờ là, những dữ liệu này cũng cho thấy miền C2 của người không hoặc chỉ tương tác yếu với L-α-phosphatidylserin như được thể hiện trong kết quả cho FP270 (EGF-HSA-C2; SEQ ID NO: 36), mà cùng với FP250 (EGF-HSA; SEQ ID NO: 32) không cạnh tranh ở định dạng thử nghiệm này. FP100, protein EGF-C2-C2 (SEQ ID NO: 26) được thử nghiệm và không cạnh tranh trong định dạng thử nghiệm này (không được thể hiện), để lại miền C1 là gốc liên kết PS chính trong MFG-E8 người. Phát hiện này gây ngạc nhiên vì phần chính của tài liệu cho thấy miền C2 của MFG-E8 là miền chính chịu trách nhiệm cho việc liên kết PS (Andersen *et al.,* (2000) Biochemistry, 39(20): 6200-6; Shi & Gilbert (2003) Blood, 101: 2628-2636; Shao *et al.,* (2008) J Biol Chem., 283(11): 7230-41). Kết luận, những phát hiện này chứng minh rằng miền C1 là miền liên kết PS tích hợp chính của các protein được thiết kế di truyền MFG-E8 và quan trọng đối với các chức năng phụ thuộc vào việc liên kết PS. Do đó, miền C1 có thể hữu ích để thay thế thành các protein khác loại để tạo ra sự liên kết PS; tuy nhiên, sự liên kết PS cao nhất đã được hiển thị đối với các protein dung hợp chứa miền nối tiếp C1-C2 hoặc C1-C1 (sau này không được hiển thị).

3.2 *Thử nghiệm bám dính Integrin αv*

Các protein dung hợp được pha loãng trong dung dịch muối đệm phosphat (PBS) độ pH 7,4 và 50µL của dung dịch 24nM được cố định bằng cách hấp phụ (đĩa 96 giếng, Nunc Maxisorb) qua đêm (1,2nM/giếng). Các đĩa này sau đó được xử lý bằng PBS chứa 3% albumin huyết thanh bò (BSA) không chứa axit béo ở nhiệt độ phòng trong 1,5 giờ. Tế bào u lympho biểu hiện integrin αvβ3 (ATCC-TIB-48 BW5147.G.1.4, ATCC, Mỹ) được nuôi cấy trong RPMI 1640 bổ sung GlutaMax, HEPES 25 mM, FBS 10%, Pen/Strep, NaPyruvate 1 mM, β -Mercaptoetanol 50 µM. Các tế bào được tách vào ngày trước thí nghiệm bám dính. Các tế bào được đánh dấu bằng 3 µg/mL 2',7'-bis-(2-carboxyetyl)-5-(và-6)-carboxyfluoresxein, axetoxymetyl este (BCECF AM) (Thermo Fisher Scientific Inc, Mỹ) trong 30 phút. Tế bào BW5147.G.1.4 được tái tạo huyền phù

trong dung dịch đệm bám dính (TBS, BSA 0,5%, 1 mM $MnCl_2$, độ pH 7,4) và 50000 tế bào/giếng được phép bám dính ở nhiệt độ phòng trong 40 phút. Các tế bào không bám dính được loại bỏ bằng cách rửa nhiều lần với đệm bám dính. Sự phát huỳnh quang của các tế bào bám dính được định lượng sử dụng thiết bị đọc đĩa đa nhãn Envision™2103, Perkin Elmer, Mỹ. Phân tích dữ liệu được thực hiện bằng phần mềm MS Excel và GraphPad Prism.

Sự bám dính tế bào với protein dung hợp cố định FP330 (EGF-HSA-C1-C2; SEQ ID NO: 42) bị phong bế hoàn toàn bởi chất ức chế integrin αv là cilengitid hoặc 10 mM EDTA chứng tỏ sự bám dính tế bào phụ thuộc vào integrin với protein được thiết kế được cố định (Hình 5A). Đột biến điểm đơn ở môtip liên kết integrin RGD (RGD > RGE) của miền giống EGF (FP280; SEQ ID NO: 38) dẫn đến việc loại bỏ hoàn toàn sự bám dính tế bào chứng tỏ rằng môtip liên kết RGD có chức năng và có thể tiếp cận được trong protein dung hợp là cần thiết cho sự bám dính phụ thuộc vào integrin αv (Hình 5B). Protein EGF-HSA cố định thiếu các miền C1-C2, FP250 (SEQ ID NO: 32) không, hoặc chỉ hỗ trợ một chút, việc bám dính của các tế bào BW5147.G.1.4 mặc dù có miền giống EGF (Hình 5C). Phát hiện này cho thấy rằng trong các điều kiện thực nghiệm được kiểm tra, vòng lặp RGD trong miền giống EGF được dung hợp với HSA có thể không đủ khả năng tiếp cận với các integrin bề mặt tế bào có thể do các lý do không gian. Sự nhiễu loạn này không rõ ràng khi C1, C2 hoặc C1-C2 được dung hợp với EGF-HSA ở vị trí đầu cùng C. Các protein tái tổ hợp theo sáng chế, ví dụ, FP330 thúc đẩy sự bám dính tế bào phụ thuộc vào integrin αv tương tự như wtMFG-E8 nếu được biểu hiện trong các tế bào CHO hoặc các tế bào HEK (Hình 5D).

Kết hợp lại với nhau, những dữ liệu này chứng minh rằng các protein dung hợp theo sáng chế liên kết với các integrin tế bào, hỗ trợ sự bám dính tế bào phụ thuộc vào integrin và chỉ ra rằng trong các protein có chèn miền HSA, miền giống EGF đầu cùng C có thể sinh lợi về mặt chức năng từ miền protein được dung hợp đầu cùng C để hỗ trợ việc liên kết integrin.

3.3 Thử nghiệm tiêu hủy tế bào chết của bạch cầu trung tính bởi đại thực bào ở người

Các tế bào đơn nhân trong máu ngoại vi của người (PBMC) được phân lập từ lớp váng máu bằng phương pháp ly tâm gradien Ficoll (Ficoll®-Paque PLUS, GE Healthcare, Thụy Điển) sau đó chọn lọc âm tính các tế bào đơn nhân bằng cách sử dụng

bộ dụng cụ phân lập Stemcell (Stemcell 19059, Vancouver, Canada). Các bạch cầu đơn nhân được biệt hóa thành đại thực bào "M0" sử dụng M-CSF tái tổ hợp của người 40 ng/mL (Yếu tố kích thích lạc khuẩn đại thực bào, R&D Systems, Mỹ) trong RPMI 1640 chứa HEPES 25mM, FBS 10%, Pen/Strep, NaPyr 1 mM, β-Merc 50 µM trong 5 ngày. Một ngày trước khi tiêu hủy tế bào chết, các đại thực bào được đánh dấu bằng PKH26 sử dụng kit Cầu nối thuốc nhuộm huỳnh quang đỏ (Sigma MINI26, Mỹ). Các tế bào được tái tạo huyền phù trong RPMI 1640 chứa HEPES 25mM, FBS 10%, Pen/Strep, NaPyr 1 mM, β-Merc 50 µM và được cấy vào các đĩa 96 giếng màu đen (Corning, Mỹ) ở 40000 tế bào/giếng và được phép bám dính trong 20 giờ.

Các bạch cầu trung tính: Các bạch cầu trung tính của người được phân lập từ các lớp váng máu bằng việc lắng dextran kết hợp với gradien mật độ Ficoll™ như sau: Huyết tương của lớp váng máu được loại bỏ bằng cách ly tâm lớp váng máu được pha loãng. Thu hoạch tế bào được pha loãng trong dextran 1% (từ Leuconostoc spp. MW 450000-650000; Sigma, Mỹ) và để lắng trên đá trong 20-30 phút.

Bạch cầu từ dịch nổi được thu hoạch và trên lớp Ficoll™-Paque (GE Healthcare Thụy Điển). Sau khi ly tâm viên nén được thu hoạch và các hồng cầu còn lại được ly giải bằng cách sử dụng đệm ly giải tế bào hồng cầu (RBC) (BioConcept, Thụy Sĩ). Các bạch cầu trung tính được rửa một lần trong môi trường (RPMI 1640 + GlutaMax chứa HEPES 25mM, FBS 10%, Pen/Strep, NaPyr 1 mM, β-Merc 50 µM) và giữ qua đêm ở 15°C. Quá trình chết theo chương trình/chết tế bào được gây ra bằng việc xử lý bạch cầu trung tính với Superfas Ligand 1 µg/mL (Enzo Life Sciences, Lausanne, Thụy Sĩ) ở 37°C trong 3 giờ. Các bạch cầu trung tính được nhuộm bằng cả Hoechst 33342 (Life Technology, Mỹ) trong 25 phút và với DRAQ5 (cBioscience, Anh, pha loãng 1:2000) ở 37°C trong tối trong 5 phút.

Thử nghiệm tiêu hủy tế bào chết

Các đại thực bào M0 được ủ với các protein dung hợp trong 30 phút. Các bạch cầu trung tính được đánh dấu chết theo chương trình được thêm vào với tỷ lệ M0/bạch cầu trung tính 1:4. Quá trình tiêu hủy tế bào chết của các bạch cầu trung tính chết theo chương trình bởi các đại thực bào đã được làm trực quan bằng cách tận dụng sự gia tăng cường độ huỳnh quang của DRAQ5 khi định vị các bạch cầu trung tính trong ngăn chứa lysosom có độ pH thấp của các đại thực bào M0.

Sự tiêu hủy tế bào chết được định lượng sử dụng hệ thống phân tích hàm lượng cao trường rộng ImageXpress Micro XLS (Molecular DEVICES. CA, Mỹ). Các đại thực bào được xác định thông qua huỳnh quang PKH26. Quá trình tiêu hủy tế bào chết (EI, hiển thị dưới dạng %) được tính bằng tỷ lệ giữa các đại thực bào có chứa ít nhất một sự kiện bạch cầu trung tính chết theo chương trình (DRAQ5high) được tiêu hóa trên tổng số các đại thực bào. Phân tích dữ liệu được thực hiện bằng phần mềm MS Excel và GraphPad Prism.

Tác động của protein dung hợp FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) đối với sự thúc đẩy việc tiêu hủy tế bào chết của các bạch cầu trung tính sắp chết bởi các đại thực bào ở người được thể hiện trên Hình 6. Các protein dung hợp làm tăng quá trình nhập nội của các bạch cầu trung tính sắp chết ở người được đánh dấu bằng pHrodo vào đại thực bào với khả năng tiêu hủy tế bào chết vốn đã cao của các đại thực bào M0, được thể hiện ở mức cơ bản. Hình 7 thể hiện rằng protein dung hợp tái tổ hợp FP278 có thể giải cứu sự tiêu hủy tế bào chết bị suy giảm bởi lipopolysacarit) nội độc tố của các bạch cầu trung tính sắp chết bởi các đại thực bào ở người. Hình 7A thể hiện sự suy giảm của quá trình tiêu hủy tế bào chết bởi các đại thực bào của các bạch cầu trung tính sắp chết của người là bằng 100 pg/ml lipopolysaccarit (LPS) ở ba đối tượng cho là người. Bảng bên trái thể hiện sự đáp ứng đối tượng cho riêng rẽ, bảng bên phải thể hiện sự suy yếu trung bình của sự tiêu hủy tế bào chết (%) của ba đối tượng cho. Hình 7B thể hiện sự giải cứu quá trình tiêu hủy tế bào chết bị suy giảm bởi nội độc tố (LPS) này của các bạch cầu trung tính sắp chết bởi các đại thực bào ở người với protein dung hợp FP278.

Sự giải cứu quá trình tiêu hủy tế bào chết bị suy giảm bởi hạt *S. aureus* của bạch cầu trung tính sắp chết bởi đại thực bào người với protein dung hợp FP330 được thể hiện trên Hình 8. Hình 8A thể hiện tác động của nồng độ 100 nM của protein dung hợp trong việc thúc đẩy sự tiêu hủy tế bào chết ở mức cơ bản (đường chấm; phần bên trái của hình) cũng như tác dụng của protein dung hợp 100 nM trong việc giải cứu sự suy giảm của việc tiêu hủy tế bào chết được gây ra bởi việc bổ sung *S. aureus* (phần bên phải của hình). Hình 8B thể hiện tác động của các nồng độ tăng dần của protein dung hợp FP278 (EC_{50} 8nM) đối với việc giải cứu sự tiêu hủy tế bào chết bị suy giảm do việc bổ sung của *S. aureus* gây ra, và đối với việc thúc đẩy sự tiêu hủy tế bào chết khi các mức cơ sở của sự tiêu hủy tế bào chết đã đạt được.

3.4 Thử nghiệm tiêu hủy tế bào chết nội mô – Jurkat người

Nuôi cấy tế bào

Tế bào nội mô tĩnh mạch rốn người (HUVEC) thu được từ Lonza (Basel, Thụy Sĩ). Các tế bào được nuôi cấy trong các bình có phủ gelatin (từ da bò, nồng độ cuối cùng trong PBS là 0,2%, dung dịch pha loãng của dung dịch gốc 2%, Sigma, Đức). Các tế bào được nuôi cấy với môi trường nuôi cấy 199 (Thermo Fischer Scientific, Mỹ) bổ sung FBS 10% (GE Healthcare, Vương quốc Anh), Pen/Strep 1% (Thermo Fischer Scientific, Mỹ), Glutamax 1% (Thermo Fischer Scientific, Mỹ) và Yếu tố sinh trưởng nguyên bào sợi tái tổ hợp - cơ bản 1 ng/mL (Peprotech, Vương quốc Anh). Các tế bào được tách ra để thu hoạch hoặc cấy chuyển sử dụng Accutase™ (Thermo Fischer Scientific, Mỹ).

Các tế bào Jurkat E6-1 thu được từ ATCC (Bộ sưu chủng chuẩn Hoa Kỳ, Mỹ) và được nuôi trong môi trường nuôi cấy RPMI 1640 (Thermo Fischer Scientific, Mỹ) được bổ sung FBS 10% (GE Healthcare, Vương quốc Anh), Pen/Strep 1% (Thermo Fischer Scientific, Mỹ), Natri Pyruvat 10 mM (Thermo Fischer Scientific, Mỹ) và HEPES 10 mM axit (4-(2-hydroxyetyl)-1-piperazinetansulfonic, Thermo Fischer Scientific, Mỹ).

Sự chết theo chương trình của các tế bào Jurkat E6-1 được tạo ra sử dụng TRAIL tái tổ hợp ở người (R&D Systems, Mỹ). Tế bào chết theo chương trình được đánh dấu bằng thuốc nhuộm pHrodo™ xanh lá cây STP este (Thermo Fischer Scientific, Mỹ). Dung dịch đệm đếm tế bào theo dòng được chuẩn bị với PBS (Thermo Fischer Scientific, Mỹ) bổ sung FBS 1% (GE Healthcare, Vương quốc Anh), 0,05% khối lượng/thể tích natri azit (Merck, Đức) và EDTA 0,5 mM (axit Etylendiaminetetraaxetic, Thermo Fischer Scientific, Mỹ).

Thử nghiệm tiêu hủy tế bào chết

Vào ngày 1, các HUVEC (hợp lưu 70-90%) được thu hoạch bằng cách tách riêng với Accutase™ trong 5 phút được rửa bằng PBS và được tái tạo huyền phù trong môi trường nuôi cấy tế bào. Số lượng và khả năng sống của tế bào được đánh giá bằng cách sử dụng máy đếm tế bào theo dòng Guava EasyCyte (Merck, Đức) và chất phản ứng Guava ViaCount (Merck, Đức) theo hướng dẫn của nhà sản xuất. Lượng tế bào cần thiết được ly tâm ở 300xg trong 5 phút ở nhiệt độ phòng và được tái tạo huyền phù trong môi trường nuôi cấy để cho phép số lượng tế bào là $6,6x10^4$ tế bào/mL. 150 µL/giếng của

huyền phù tế bào này được thêm vào các đĩa nuôi cấy mô 96 giếng (Corning™, Mỹ). Các HUVEC được ủ trong tủ ủ ở 37°C/5% CO_2/độ ẩm 95% trong thêm 16-20 giờ.

Số lượng tế bào Jurkat E6-1 và tình trạng khả năng sống/tế bào chết được đánh giá sử dụng máy đo tế bào theo dòng Guava EasyCyte (Merck, Đức) và chất phản ứng Guava ViaCount (Merck, Đức) theo hướng dẫn của nhà sản xuất. Lượng tế bào cần thiết được ly tâm ở 300xg trong 5 phút ở nhiệt độ trong phòng và được tái tạo huyền phù ở mật độ $1x10^6$ tế bào/mL trong môi trường nuôi cấy bổ sung TRAIL tái tổ hợp của người ở nồng độ cuối cùng là 50 ng/mL. Sự chết tế bào được gây ra ở 37°C/CO2 5%/độ ẩm 95% qua đêm.

Vào ngày thứ 2, môi trường được loại bỏ khỏi các HUVEC bằng cách hút và thêm 25 µL môi trường nuôi cấy tươi đã được làm ấm trước (37°C), tiếp theo là bổ sung 25 µL protein dung hợp hoặc các đối chứng được pha loãng trong môi trường nuôi cấy đã được làm ấm trước (37°C). Để pha loãng, các đĩa 96 giếng được xử lý bề mặt không liên kết (NBS) (Corning™, Mỹ) được sử dụng. Các protein dung hợp được cho tương tác với các HUVEC trong 30 phút ở 37°C/CO_2 5%/độ ẩm 95% trước khi bổ sung các tế bào Jurkat sắp chết.

Số lượng tế bào Jurkat E6-1 chết theo chương trình/sắp chết được đếm sử dụng máy đo tế bào theo dòng Guava EasyCyte (Merck, Đức) và thuốc thử Guava ViaCount (Merck, Đức). Lượng tế bào chết theo chương trình cần thiết được ly tâm ở 400xg ở nhiệt độ phòng trong 5 phút và được tái tạo huyền phù ở mật độ $5x10^6$ tế bào/mL trong môi trường RPMI 1640 (không FBS) được bổ sung với thuốc nhuộm este pHrodo™ STP xanh lá cây ở nồng độ cuối cùng là 5 µg/mL (Môi trường nhuộm). Sau khi nhuộm trong 10 phút ở 37°C, este pHrodo™ STP xanh lá cây phản ứng còn lại được bất hoạt với môi trường nhuộm được bổ sung FBS 10% trong 5 phút nữa ở 37°C. Các tế bào được đánh dấu pHrodo™ xanh lá cây được rửa một lần và số lượng tế bào được điều chỉnh thành $3x10^6$ tế bào/mL trong môi trường nuôi cấy HUVEC. $1,5x10^6$/giếng các tế bào Jurkat được đánh dấu pHrodo™ xanh lá cây được thêm vào các HUVEC và được ủ ở 37°C/CO_2 5%/độ ẩm 95% trong 5 giờ. Môi trường đã được loại bỏ, các HUVEC được rửa một lần trong PBS và được tách bằng 40 µL/giếng dung dịch Accutase™. Tế bào được thu hoạch bằng cách bổ sung 80 µL đệm đếm tế bào theo dòng nước đá lạnh, được chuyển vào 1,5 mL khối 96 giếng polypropylen, được rửa sạch bằng lượng dư đệm đếm tế bào theo dòng nước đá lạnh và ly tâm ở 400xg (4°C) trong 5 phút. Các dịch nổi được

loại bỏ bằng cách hút và các hạt được tái tạo huyền phù trong dung dịch đệm đếm tế bào theo dòng nước đá lạnh 80 µL và được chuyển vào đĩa vi chuẩn độ đáy V 96 giếng (BD Biosciences, Mỹ). Sau đó, các mẫu được đo trên máy đếm tế bào theo dòng BD LSRFortessa™ (BD Biosciences, Mỹ). Cường độ huỳnh quang xanh lá pHrodo™, là chỉ thị của sự định vị lysosom của các tế bào Jurkat được nhận chìm, được ghi lại. Phân tích dữ liệu đếm tế bào theo dòng được thực hiện bằng cách sử dụng phần mềm FlowJo™. Các giá trị cường độ huỳnh quang trung bình (MFI) của tín hiệu pHrodo™ xanh lá cây từ các HUVEC cổng đơn lẻ được sử dụng làm kết quả đọc. Phân tích dữ liệu được thực hiện sử dụng phần mềm MS Excel và GraphPad Prism để tính toán EC_{50}.

Tác động của protein dung hợp FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) và FP270 (EGF-HSA-C2; SEQ ID NO: 36) lên sự thúc đẩy sự tiêu hủy tế bào chết của các tế bào Jurkat sắp chết bởi các tế bào nội mô HUVEC được thể hiện trên Hình 9. Việc nội hóa các tế bào T Jurkat sắp chết ở người được đánh dấu bằng pHrodo bởi các HUVEC được thúc đẩy mạnh mẽ bởi protein dung hợp FP278. Các kết quả chứng minh rằng các tế bào nội mô được trang bị bởi protein dung hợp để trở thành các thể thực bào hiệu quả của các tế bào sắp chết. Đáng ngạc nhiên là hiệu quả của các protein dung hợp trong thử nghiệm này rõ ràng phụ thuộc vào sự hiện diện của miền nối tiếp C1-C2 hoặc C1-C1. Protein dung hợp bao gồm EGF-HSA-C2 (FP270), ví dụ không hoạt động trong thiết đặt thử nghiệm này, như thể hiện trên Hình 9. Hình 10 chứng minh phát hiện bất ngờ của các tác giả sáng chế rằng vị trí của miền HSA trong các protein được thiết kế, cụ thể là ở vị trí đầu cùng N hoặc C (HSA-EGF-C1-C2 (FP220; SEQ ID NO: 30) hoặc EGF-C1-C2-HSA (FP110; SEQ ID NO: 28), tương ứng), đem lại khả năng phong bế sự tiêu hủy tế bào chết ở thử nghiệm tiêu hủy tế bào chết bởi đại thực bào đối với protein được thiết kế MFG-E8 HSA. Dữ liệu này chứng minh rõ ràng tầm quan trọng của việc định vị vùng HSA giữa các miền liên kết integrin và miền liên kết PS để thúc đẩy hiệu quả sự tiêu hủy tế bào chết bởi các protein dung hợp theo sáng chế.

Hình 11 thể hiện sự so sánh sự thúc đẩy việc tiêu hủy tế bào chết nội mô bởi các định dạng khác nhau của các protein dung hợp bao gồm các sự kết hợp của miền EGF, miền C1-C2, HSA hoặc miền Fc. Hình 11A thể hiện sự so sánh các protein dung hợp bao gồm HSA với HSA được định vị ở đầu cùng C hoặc đầu cùng N hoặc giữa các miền giống EGF và miền C1-C2; EGF-C1-C2-HSA (FP110; SEQ ID NO: 28), HSA-EGF-C1-C2 (FP220; SEQ ID NO: 30) và EGF-HSA-C1-C2-thẻ His (FP278; SEQ ID NO:

44), tương ứng. Hình 11B thể hiện việc so sánh các protein dung hợp bao gồm miền Fc với Fc được định vị ở đầu cùng C hoặc giữa các miền C1 và giống EGF. Hai định dạng của gốc Fc được thể hiện: Fc kiểu dại (SEQ ID NO: 7) như được tìm thấy ở FP070 (EGF-Fc-C1-C2; SEQ ID NO: 17) và FP080 (EGF-C1-C2-Fc; SEQ ID NO: 22) và các gốc Fc với các cải biến KiH S354C và T366W trên một nhánh của Fc (FP060; EGF-C1-C2-Fc [S354C, T366W]; SEQ ID NO: 14) cách đánh số châu Âu (Merchant *et al* (1998) *nêu trên*). Hình 11C thể hiện sự so sánh các protein dung hợp FP090 (Fc-EGF-C1-C2; SEQ ID NO: 24) bao gồm gốc Fc được định vị ở đầu cùng N, đối với hai mẻ FP090 ở ba nồng độ khác nhau (0,72, 7,2 và 72nM) so với đối chứng wtMFG-E8. Quá trình tiêu hủy tế bào chết của các tế bào Jurkat sắp chết bởi các HUVEC chỉ được thúc đẩy bởi các protein được thiết kế với gốc HSA hoặc Fc được chèn sau miền giống EGF. Hình 11D thể hiện sự chèn miền hòa tan có thể dẫn đến protein dung hợp hoạt tính sinh học mới dựa trên protein bắc cầu nội sinh EDIL3, đoạn cùng nguồn gốc của MFG-E8. Như được thể hiện trên Hình 11D, HSA được chèn ở giữa miền giống EGF và miền C1-C2 của EDIL3, đoạn cùng nguồn gốc của MFG-E8. Cấu trúc EDIL3 này (FP050 (EGF-HSA-C1-C2 dựa vào EDIL3; SEQ ID NO: 12) có chỉ một (chứa vòng lặp RGD) trong số 3 miền giống EGF mà được tìm thấy ở wtEDIL3. Trong cấu trúc này, các tác giả sáng chế đã bất ngờ tìm thấy sự dung nạp tương tự của việc chèn miền HSA liên quan đến sự biểu hiện của protein được thiết kế tái tổ hợp mới với độ tinh sạch rất cao (Hình 2B). Ngoài ra, điều đáng ngạc nhiên là protein được thiết kế tái tổ hợp FP050 có nguồn gốc EDIL3 thúc đẩy việc tiêu hủy tế bào chết của các tế bào Jurkat sắp chết bởi các tế bào nội mô (HUVECS) thể hiện chức năng cốt lõi của protein bắc cầu và minh chứng rằng các miền của các protein bắc cầu là hữu dụng để thiết kế các protein được thiết kế tái tổ hợp mới về chức năng.

Ví dụ 4: Quá trình tiêu hủy tế bào chết của vi hạt huyết tương tiền huyết khối

4.1 Thử nghiệm sự tiêu hủy tế bào của nội mô-vi hạt người

Nuôi cấy tế bào

Tế bào HUVEC thu được từ Lonza (Basel, Thụy Sĩ). Các tế bào được nuôi cấy trong các bình phủ gelatin (từ da bò, nồng độ cuối cùng là 0,2% trong PBS, dung dịch pha loãng của dung dịch gốc 2%, Sigma Aldrich/Merck, Đức). Các tế bào được nuôi cấy với môi trường nuôi cấy 199 (Thermo Fischer Scientific, Mỹ) bổ sung FBS 10%

(GE Healthcare, Vương quốc Anh), Pen/Strep 1% (Thermo Fischer Scientific, Mỹ), Glutamax 1% (Thermo Fischer Scientific, Mỹ) và Yếu tố sinh trưởng tế bào sợi tái tổ hợp cơ bản 1 ng/mL (Peprotech, Vương quốc Anh). Các tế bào được tách ra để thu hoạch hoặc cấy chuyển sử dụng Accutase™ (Thermo Fischer Scientific, Mỹ).

Các vi hạt có nguồn gốc từ tiểu cầu được điều chế theo quy trình sau: máu tĩnh mạch được xitrat hóa được thu thập (Coagulation 9NC Citrate Monovette, Sarstedt, Đức) từ những người tình nguyện trưởng thành khỏe mạnh sau khi được sự đồng ý bằng văn bản. Huyết tương giàu tiểu cầu (PRP) được tạo ra bằng cách ly tâm (200xg, 15 phút, không hãm, nhiệt độ phòng). Các vi hạt/mảnh vụn có nguồn gốc từ tiểu cầu được tạo ra bằng cách đưa PRP vào ba chu kỳ tách/đông lạnh bằng cách sử dụng nitơ lỏng và rã đông ở 37°C. Các mảnh/vi hạt tiểu cầu được tạo thành viên bằng cách ly tâm ở 20'000xg trong 15 phút ở nhiệt độ phòng. Viên nén được tái hòa tan trong PBS, các phần nhỏ được tạo ra và bảo quản ở -80°C. Các chế phẩm vi hạt có 85-100% PS dương tính được xác định bằng phương pháp đếm tế bào theo dòng sử dụng Alexa Fluor™ MFG-E8/lactadherin ở chuột được đánh dấu 488 (Novartis nội bộ). Số lượng vi hạt được xác định bằng cách sử dụng các hạt đếm chuyên dụng (BioCytex / Stago, Pháp). Dung dịch đệm đếm tế bào theo dòng được chuẩn bị với PBS (Thermo Fischer Scientific, Mỹ) bổ sung FBS 1% (GE Healthcare, Vương quốc Anh), 0,05% khối lượng/thể tích natri azit (Merck, Đức) và EDTA 0,5 mM (axit Etylendiaminetetraaxetic, Thermo Fischer Scientific, Mỹ).

4.2 Thử nghiệm quá trình tiêu hủy tế bào chết

Vào ngày 1, các tế bào HUVEC (hợp lưu 70-90%) được thu hoạch bằng cách tách riêng với Accutase™ trong 5 phút được rửa bằng PBS và được tái tạo huyền phù trong môi trường nuôi cấy tế bào. Số lượng và khả năng sống của tế bào được đánh giá bằng cách sử dụng máy đếm tế bào theo dòng Guava EasyCyte (Merck, Đức) và chất phản ứng Guava ViaCount (Merck, Đức) theo hướng dẫn của nhà sản xuất. Lượng tế bào cần thiết được ly tâm ở 300xg trong 5 phút ở nhiệt độ phòng và được tái tạo huyền phù trong môi trường nuôi cấy để cho phép số lượng tế bào là $6,6x10^4$ tế bào/mL. 150 µL/giếng của huyền phù tế bào này được thêm vào các đĩa nuôi cấy mô 96 giếng (Corning™, Mỹ). Các tế bào HUVEC được ủ trong thiết bị ủ ở 37°C/CO2 5%/độ ẩm 95% trong thêm 16-20 giờ.

Ở ngày thứ 2, môi trường được loại bỏ khỏi các tế bào HUVEC bằng cách hút và thêm 25 µL môi trường nuôi cấy tươi đã được làm ấm trước (37°C), tiếp theo là bổ sung 25 µL protein dung hợp FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) ở ba nồng độ khác nhau: 0,3nM, 3nM hoặc 30nM hoặc đối chứng. được pha loãng trong miễn dịch nuôi cấy được làm ấm trước (37°C). Để pha loãng, các đĩa 96 giếng được xử lý bề mặt không liên kết (NBS) (Corning™, Mỹ) được sử dụng. Các protein được thử nghiệm được cho phép tương tác với tế bào HUVEC ở 37°C / 5% CO_2 / độ ẩm 95% trong 30 phút trước khi được bổ sung các vi hạt có nguồn gốc từ tiểu cầu.

Lượng các vi hạt cần thiết được ly tâm ở 20'000xg ở 4°C trong 15 phút và được tái tạo huyền phù ở mật độ $2x10^8$ hạt/mL trong môi trường RPMI 1640 (không FBS) được bổ sung với thuốc nhuộm este pHrodo™ STP xanh lá cây ở nồng độ cuối cùng là 5 µg/mL (Môi trường nhuộm). Sau khi nhuộm trong 10 phút ở 37°C, este pHrodo™ STP xanh lá cây phản ứng còn lại được bất hoạt với môi trường nhuộm được bổ sung FBS 10% trong 5 phút nữa ở 37°C. Các vi hạt đánh dấu pHrodo™ xanh lá cây được rửa một lần bằng việc ly tâm ở 20'000xg ở 4°C trong 15 phút và số lượng tế bào được điều chỉnh thành $1x10^8$ tế bào/mL trong môi trường nuôi cấy HUVEC. $1,5x10^6$ hạt/giếng các tế bào Jurkat đánh dấu pHrodo™ xanh lá cây được thêm vào các tế bào HUVEC và được ủ ở 37°C/CO_2 5%/độ ẩm 95% trong 5 giờ. Môi trường được loại bỏ, các tế bào HUVEC được rửa một lần trong PBS và được tách bằng 40 µL/giếng dung dịch Accutase™. Tế bào được thu hoạch bằng cách bổ sung 80 µL đệm đếm tế bào theo dòng nước đá lạnh, chuyển vào khối 1,5 mL polypropylen 96 giếng, rửa sạch bằng lượng dư đệm đếm tế bào theo dòng nước đá lạnh và ly tâm ở 400xg (4°C) trong 5 phút. Các dịch nổi được loại bỏ bằng cách hút và các hạt được tái tạo huyền phù trong dung dịch đệm đo tế bào dòng nước đá lạnh 80 µL và được chuyển vào đĩa vi chuẩn độ đáy V 96 giếng (BD Biosciences, Mỹ). Các mẫu được đo trên máy đếm tế bào theo dòng BD LSRFortessa™ (BD Biosciences, Mỹ). Cường độ huỳnh quang xanh lá pHrodo™, là chỉ thị của sự định vị lysosom của các vi hạt được nhận chìm, được ghi lại. Phân tích dữ liệu đếm tế bào theo dòng được thực hiện bằng cách sử dụng phần mềm FlowJo™. Các giá trị cường độ huỳnh quang trung bình (MFI) của tín hiệu pHrodo™ xanh lá cây từ các tế bào HUVEC cổng đơn lẻ được sử dụng làm kết quả đọc. Phân tích dữ liệu được thực hiện sử dụng phần mềm MS Excel và GraphPad Prism để tính toán EC_{50}. Protein dung hợp FP278 thúc đẩy quá trình tiêu hủy tế bào chết của các vi hạt có nguồn gốc tiểu

cầu bởi các tế bào nội mô theo cách phụ thuộc nồng độ như được thể hiện trên Hình 12. Sự thúc đẩy hấp thu phụ thuộc vào nồng độ và cũng được quan sát thấy ở các loại tế bào nội mô khác (không được hiển thị).

Ví dụ 5: Các đặc tính kỹ thuật của các protein dung hợp MFG-E8-HSA

5.1 Thử nghiệm liên kết Cộng hưởng plasmon bề mặt (SPR) của protein dung hợp FP330

với FcRn

Thử nghiệm liên kết trực tiếp được thực hiện để mô tả sự liên kết của protein dung hợp FP330 (EGF-HSA-C1-C2; SEQ ID NO: 42) với FcRn. Hằng số ái lực liên kết động học (KD) được đo trên protein được bắt chụp sử dụng FcRn tái tổ hợp người làm chất phân tích. Các phép đo được tiến hành trên BIAcore® T200 (GE Healthcare, Glattbrugg, Thụy Sĩ) ở nhiệt độ phòng và độ pH tương ứng là 5,8 và 7,4. Đối với phép đo ái lực, các protein được pha loãng trong NaP 10mM, NaCl 150mM, Tween 20 0,05%, độ pH 5,8 và được cố định trên các tế bào dòng của chip cảm biến cấp nghiên cứu CM5 (GE Healthcare, ref BR-1000-14) sử dụng quy trình tiêu chuẩn theo khuyến cáo của nhà sản xuất (GE Healthcare). Để làm đối chứng tham chiếu, một tế bào dòng được cố định trống. Dữ liệu liên kết được thu thập bằng cách bơm liên tiếp hàng loạt các dung dịch pha loãng của chất phân tích vào tế bào dòng tham chiếu và đo. Các mẫu có nồng độ bằng không (chỉ chạy đệm) được đưa vào để cho phép tham chiếu kép trong quá trình đánh giá dữ liệu. Để đánh giá dữ liệu, các biểu đồ cảm biến được tham chiếu kép được sử dụng và các hằng số phân ly (KD) được phân tích.

Protein dung hợp FP330 liên kết với FcRn ở độ pH 5,8 với ái lực 1380nM, trong khi không có sự liên kết nào được quan sát ở độ pH 7,4 (Xem Bảng 5 ở trên). Các kết quả này phù hợp tốt với HSA kiểu dại (1000-2000 nM, ở độ pH 5,8, dữ liệu không được hiển thị).

5.2 Đo nhiệt lượng quét vi sai (DSC) của MFG-E8 và các biến thể

Độ ổn định nhiệt của biến thể protein MFG-E8 được thiết kế FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) được đo bằng cách sử dụng phép đo nhiệt lượng quét vi sai. Các phép đo được thực hiện trên nhiệt lượng kế vi mô quét vi sai (Nano DSC, TA instruments). Thể tích tế bào là 0,5ml và tốc độ gia nhiệt là 1°C/phút. Protein được sử

dụng ở nồng độ 1mg/ml trong PBS (độ pH 7,4). Nhiệt dung mol của protein được ước tính bằng cách so sánh với các mẫu trùng lặp có chứa đệm giống hệt nhau mà từ đó protein đã bị loại bỏ. Nhiệt dung từng phần mol và đường cong nóng chảy được phân tích sử dụng quy trình tiêu chuẩn. Các biểu đồ nhiệt đã được hiệu chỉnh đường cơ sở và nồng độ được chuẩn hóa. Hai sự kiện nóng chảy đã được quan sát, Tm đầu tiên ở 50°C, Tm thứ hai ở 64°C.

5.3 Các phép đo khuynh hướng kết tụ và độ hòa tan của các biến thể MFG-E8

Thứ nhất, xu hướng kết tụ của protein biến thể MFG-E8 FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID NO: 44) được đo bằng sự tán xạ ánh sáng động (DLS, Wyatt). Tán xạ ánh sáng động được áp dụng để đo các hệ số khuếch tán tịnh tiến của FP278 trong dung dịch bằng cách định lượng các dao động động của ánh sáng tán xạ. Sự phân bố kích thước biến thể protein mà không phân đoạn, cung cấp các ước tính độ phân tán cũng như bán kính thủy động lực học được đo ở nồng độ 1mg/ml. Bán kính thủy động lực học của protein dung hợp FP278 được xác định bằng đầu đọc đĩa DynaPro ™ (Wyatt Technology Europe GmbH, Dernbach, Đức) kết hợp với phần mềm DYNAMICS (phiên bản 7.1.0.25, Wyatt). 50 µL dung dịch protein không pha loãng và được lọc (Bộ lọc PVDF 0,22 µm (Millex® Syringe-driven Filter Unit, Millipore, Billerica, Mỹ)) được đo trong đĩa 384 giếng (đĩa giếng tròn 384, Polystyrol, Thermo Scientific, Langenselbold, Đức). Không thể xác định được các khối tụ có khối lượng phân tử cao hơn của mẫu protein. Bán kính thủy động lực học của protein là khoảng 5-6nm, cho thấy protein monome trong dung dịch.

Thứ hai, các phép đo bán kính thủy động lực học phụ thuộc nồng độ của protein dung hợp FP278 được thực hiện để ước tính độ hòa tan của protein. Các nồng độ protein lên đến 22 mg/ml được áp dụng. Bán kính thủy động lực học được xác định như mô tả ở trên. Khi tăng nồng độ của protein dung hợp FP278, không thể quan sát thấy sự gia tăng bán kính (5-7 nm), trong khi phép đo tán xạ ánh sáng động của wtMFG-E8 (SEQ ID NO: 1) không thành công do sự kết tụ cao ở nồng độ khoảng 0,2mg/ml.

Ví dụ 6: Tối ưu hóa các protein dung hợp MFG-E8

Khối phổ (MS) được sử dụng để nghiên cứu protein dung hợp FP330 (EGF-HSA-C1-C2) để tạo ra bảng các protein dung hợp dựa trên MFG-E8 biến thể được tối ưu hóa để cải thiện sự biểu hiện và năng suất. Bảng các protein biến thể được tạo ra với các cầu

nối có kích thước và cấu trúc khác nhau, ví dụ, các cầu nối bao gồm GS giữa các miền EGF và HSA và/hoặc bội số của GS hoặc G4S (SEQ ID NO: 64) giữa các miền HSA và C1. Ngoài ra, các cải biến amino axit (được mô tả là HSA* trong Bảng 7) bao gồm đột biến mất đoạn hoặc đột biến thế được bao gồm trong một số biến thể. Bảng các protein dung hợp biến thể được tổng hợp trong Bảng 7 dưới đây.

Bảng 7: Tổng hợp các protein dung hợp biến thể

Biến thể	**Miền**	**Sự cải biến axit amin[1]**	**Cầu nối**	**SEQ ID NO:**
wtMFG-E8	EGF-C1-C2	-	-	1
FP330	EGF-GS-HSA-cầu nối-C1-C2	-	$(G_2S)_4$ cầu nối (SEQ ID 62)	42
FP278	EGF-GS-HSA-cầu nối-C1-C2-thẻ His	-	$(G_2S)_4$ cầu nối (SEQ ID 62)	44
FP811	EGF-GS-HSA*-cầu nối-C1-C2	Sự làm khuyết: G632 - L633	G_4S (SEQ ID NO: 64)	54
FP010	EGF-GS-HSA*-cầu nối-C1-C2	Sự làm khuyết: G632 - L633	$(G_4S)_2$ (SEQ ID NO: 65)	56
FP816	EGF-HSA-C1-C2	-	-	58
FP138	EGF-GS-HSA*-cầu nối-C1-C2	Sự làm khuyết: G632 - L633	$(G_2S)_4$ cầu nối (SEQ ID 62)	52

FP284	EGF-GS-HSA*-cầu nối-C1-C2	Sự thế L633V	$(G_2S)_4$ cầu nối (SEQ ID 62)	50
FP776	EGF-HSA*-C1-C2	Sự làm khuyết: A626 - L633	-	48
FP068	EGF-HSA*-C1-C2	Sự làm khuyết: G632 - L633	-	46

[1] Vị trí của sự cải biến axit amin được đánh số theo SEQ ID NO: 42 (FP330)

Ví dụ 7: Các protein dung hợp MFG-E8 biến thể; biểu hiện và tinh chế

Các phương pháp tạo ra các protein dung hợp trong các dòng tế bào HEK được mô tả trong Ví dụ 2. Để biểu hiện trong dòng tế bào CHO, các axit nucleic mã hóa các biến thể MFG-E8 đã được tổng hợp tại Geneart (LifeTechnologies) và được tách dòng vào trong vectơ biểu hiện động vật có vú bằng cách sử dụng các kỹ thuật tách dòng dựa trên sự ghép nối bằng enzym cắt giới hạn. Các plasmit thu được được chuyển nhiễm vào các tế bào CHO-S (Thermo). Nói tóm lại, đối với việc biểu hiện thoáng qua của các protein dung hợp, vectơ biểu hiện được chuyển nhiễm vào các tế bào CHO-S được thích ứng huyền phù sử dụng tác nhân chuyển nhiễm ExpifectamineCHO (Thermo). Thông thường, 400 ml tế bào ở dạng huyền phù với mật độ 6 triệu tế bào mỗi ml được biến nạp ADN chứa 400 µg vectơ biểu hiện mã hóa các chuỗi được thiết kế di truyền. Sau đó, vectơ biểu hiện tái tổ hợp được đưa vào các tế bào chủ để tiết thêm trong bảy ngày trong môi trường nuôi cấy (môi trường biểu hiện ExpiCHO, được bổ sung thức ăn ExpiCHO và thuốc thử tăng cường (Thermo)).

Như có thể thấy từ dữ liệu biểu hiện trong Bảng 8, các protein dung hợp biến thể FP068 (SEQ ID NO: 46) và FP776 (SEQ ID NO: 48) cho thấy cải thiện sự biểu hiện xấp xỉ hai lần so với protein dung hợp FP330 (SEQ ID NO: 42).

Bảng 8: Sự biểu hiện của các protein dung hợp biến thể trong các dòng tế bào HEK và CHO*

Protein	**Biểu hiện sau bắt giữ HSA (mg/l)**
FP330	11
FP138	10
FP816	9
FP068*	18
FP776*	21
FP284	10
FP811	8
FP010	10

* cho biết protein dung hợp được sản xuất ở dòng tế bào CHO

Ví dụ 8: Đặc tính của các protein dung hợp biến thể

Ảnh hưởng của các protein dung hợp biến thể đối với việc tiêu hủy tế bào chết được xác định bằng việc thực hiện các thử nghiệm tiêu hủy tế bào chết như được mô tả trong Ví dụ 3.

Trong thử nghiệm đầu tiên, ảnh hưởng của các protein dung hợp biến thể trong thử nghiệm tiêu hủy tế bào chết bởi đại thực bào-bạch cầu trung tính ở người được xác định theo phương pháp được mô tả trong Phần 3.3 ở trên. Các đại thực bào M0 được ủ với protein dung hợp FP330 (EGF-HSA-C1-C2; SEQ ID No: 42) hoặc các biến thể FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID No: 44) hoặc FP776 (EGF-HSA-C1-C2; SEQ ID No: 48) trong 30 phút. Như thể hiện trên Hình 13, các protein dung hợp FP330, FP278 và FP776 có thể giải cứu endotoxin (lipopolysacarit (LPS)) - làm suy giảm sự tiêu hủy tế bào chết của các bạch cầu trung tính sắp chết bởi các đại thực bào ở người. Các nồng độ tăng dần của các protein dung hợp FP330 (EC_{50} = 1,6 nM; Hình 13A),

FP278 (EC_{50} = 1,78 nM; Hình 13B) và FP776 (EC_{50} = 0,5 nM; Hình 13C) dẫn đến sự giải cứu quá trình tiêu hủy tế bào chết bị suy giảm do bổ sung LPS và thậm chí thúc đẩy việc tiêu hủy tế bào chết khi đã đạt đến các mức cơ bản.

Các protein dung hợp FP330, FP278 và FP776 được tìm hiểu đặc tính thêm trong thử nghiệm tiêu hủy tế bào chết của tế bào Jurkat bởi tế bào nội mô (HUVEC) ở người theo phương pháp được mô tả trong Phần 3.4 ở trên. Tác động của các protein dung hợp FP330, FP278 và FP776 lên sự thúc đẩy việc tiêu hủy tế bào chết của các tế bào Jurkat sắp chết bởi các tế bào nội mô HUVEC được thể hiện trên Hình 14. Việc nội hóa của của các tế bào T Jurkat sắp chết ở người được đánh dấu bằng pHrodo bởi các HUVEC bằng cách tăng đáng kể nồng độ FP330 (EC_{50} = 3,4 nM; Hình 14A), FP278 (EC_{50} = 2,4 nM; Hình 14B) và FP776 (EC_{50} = 3 nM; Hình 14C). Những kết quả này chứng minh rằng các tế bào nội mô được trang bị các protein dung hợp để trở thành các thể thực bào hiệu quả của các tế bào sắp chết.

Ví dụ 9: Sự bảo vệ chuột nhắt khỏi AKI và sự đáp ứng cơ quan cấp tính được kích hoạt bởi AKI

9.1 Mô hình tổn thương thận cấp

Chuột cái C57BL/6 (18-22 g) được mua từ Charles River (Pháp) và được nuôi nhốt trong khoang kiểm soát nhiệt độ trong các lồng được bảo vệ phía trên bằng bộ lọc với các chu kỳ sáng/tối 12 giờ. Động vật được xử lý tuân thủ nghiêm ngặt luật liên bang Thụy Sĩ và Nguyên tắc NIH về Chăm sóc Động vật trong Phòng thí nghiệm. Protein dung hợp trị liệu được thử nghiệm được sử dụng trong màng bụng (i.p.) hoặc trong tĩnh mạch (i.v.) hai giờ trước khi phẫu thuật. Buprenorphin (Indivior Schweiz AG) được dùng dưới da (s.c.) với liều 0,1 mg/kg 60 đến 30 phút trước khi phẫu thuật. Gây mê qua đường hô hấp với isofluran được thực hiện trong buồng gây mê (3,5-5% thể tích, khí mang: oxy) trong 5 phút trước khi phẫu thuật. Trong quá trình phẫu thuật, con vật được duy trì gây mê qua mặt nạ với 1-2 Vol% isofluran/oxy, tốc độ dòng khí 0,8- 1,2 l/phút. Da bụng được cạo và khử trùng bằng Betas trùng (Mundipharma, Pháp). Các con vật được đặt trên tấm chăn thu nhiệt (Rothacher- Thụy Sĩ) có hệ thống theo dõi nhiệt độ đồng căn (PhysiTemp, Mỹ- Physitemp Instruments LLC, Mỹ) và được bao phủ bởi gạc vô trùng. Nhiệt độ cơ thể được theo dõi trong suốt phẫu thuật bởi đầu dò trực tràng (Physitemp Instruments LLC, US) và được kiểm soát để cho phép nhiệt độ cơ thể là

36,5-37,5°C. Tất cả các con vật bao gồm đối chứng Giả bộ trải qua phẫu thuật cắt bỏ thận một bên của thận phải: Sau khi rạch / mổ bụng giữa đường giữa, khối u trong ổ bụng được rút về bên trái để lộ thận phải. Niệu quản phải và mạch máu thận bị đứt rời và thắt lại, thận phải sau đó được cắt bỏ. Đối với động vật trải qua AKI, phần bụng được định vị ở bên phải trên gạc vô trùng và động mạch thận và tĩnh mạch bên trái được mổ để cho phép kẹp để gây thiếu máu cục bộ. Kẹp vi phình động mạch (B Braun, Thụy Sĩ) được sử dụng để kẹp cuống thận (động mạch và tĩnh mạch với nhau bằng một chiếc kẹp) để chặn dòng máu đến thận và gây thiếu máu cục bộ ở thận. Thiếu máu cục bộ thành công được xác nhận do sự thay đổi màu sắc của thận từ đỏ sang tím sẫm, xảy ra trong vài giây. Sau khi gây thiếu máu cục bộ (35-38 phút), kẹp vi phình động mạch được lấy ra. Nước muối ấm vô trùng (~ 2ml, 37°C) được sử dụng để rửa các chất chứa trong ổ bụng để bù nước cho các mô trước khi đóng vết thương. Sau khi rửa, 1 ml nước muối vô trùng được thêm i.p. như thay thế chất lỏng. Khi bắt đầu tái tưới máu, vết thương được đóng thành hai lớp (cơ và da, riêng biệt). Các con vật sau đó được duy trì dưới ánh đèn ấm đỏ cho đến khi hồi phục hoàn toàn. Buprenorphin được dùng lại 1 giờ và 4 giờ sau phẫu thuật với liều 0,1 mg/kg và cũng được pha vào nước uống (9,091 µg/mL). Sau 24 giờ các con vật được cho chết để phân tích.

9.2 Sử dụng các protein dung hợp trị liệu

Các protein dung hợp trị liệu FP330 (EGF-HSA-C1-C2; SEQ ID No: 42), FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID No: 44) và FP776 (EGF-HSA-C1-C2; SEQ ID No: 48) được thử nghiệm trong mô hình AKI như được mô tả ở trên với các liều được nêu trong Bảng 9 dưới đây. Đối với các nghiên cứu để phát hiện các chỉ thị huyết thanh và biểu hiện của chỉ thị qPCR, protein dung hợp FP278 được sử dụng 2 giờ trước khi phẫu thuật. FP330 và FP776 được dùng liều i.v. 30 phút trước khi khởi phát tổn thương tái tưới máu sau thiếu máu cục bộ. Đối với nghiên cứu để đo sự hấp thu chất cản quang bằng việc chụp cộng hưởng từ, protein dung hợp FP776 được định lượng dự phòng 30 phút trước khi cảm ứng AKI ở mức 1,26 mg/kg hoặc được định lượng điều trị 5 giờ sau khi gây tổn thương tái tưới máu sau thiếu máu cục bộ ở 2 mg/kg i.v.

Bảng 9: Định liều các protein dung hợp trị liệu

Protein dung hợp	**Liều (mg/kg)**	**Đường sử dụng**
FP278	0,16	i.p.
	0,50	
FP330	0,20	i.v.
	0,50	
	1,50	
FP776	0,20	i.v.
	0,75	
	1,26	
	2,00	

9.3 Kết quả đọc/Phân tích đối với sự bảo vệ AKI:

Chỉ thị huyết thanh:

Các mẫu huyết thanh được lấy trong 24 giờ sau cảm ứng tái tưới máu sau thiếu máu cục bộ và phân tích hàm lượng creatinin huyết thanh và nitơ urê máu (BUN) bằng máy phân tích phòng khám Hitachi M40 theo hướng dẫn của nhà sản xuất (Axonlab, Thụy Sĩ).

Sự biểu hiện chỉ thị qPCR ở các cơ quan:

Các cơ quan (thận, gan, phổi và tim) được thu hoạch 24 giờ sau khi cảm ứng AKI và được cắt thành từng đoạn 1 cm và được bảo quản trong đệm ARN sau (Thermo Fisher Scientific Inc, Mỹ) ở 4°C qua đêm. Các mảnh cơ quan được chuyển vào đệm RLT (RNeasy Mini Kit, Qiagen, Đức) chứa 134mM Beta-mercaptoetanol (Merck, Đức) trong các ống dung giải Matrix D (MP Biomedicals FR) và đồng nhất bằng Dụng cụ FastPrep-24 (MP Biomedicals). Mô xơ tim sau đó được phân giải bằng proteinaza K (RNeasy Mini Kit), trong khi dịch lọc thận, gan và phổi được ly tâm trực tiếp trong 3 phút ở tốc

độ tối đa trong máy ly tâm siêu nhỏ (Eppendorf, Đức). Dịch nổi được chuyển vào cột quay QIAshredder (Qiagen, Đức) và được ly tâm trong 2 phút. Việc tách chiết ARN của các dòng chảy được thực hiện theo hướng dẫn sử dụng RNeasy Mini Kit, bao gồm cả quá trình phân hủy ADNaza. Nồng độ ARN được đo bằng thiết bị Nano Drop 1000 (Thermo Fisher Scientific Inc). 2µg ARN mỗi mẫu được phiên mã ngược theo hướng dẫn sử dụng kit phiên mã ngược cADN dung lượng cao (Thermo Fisher Scientific Inc) bằng cách sử dụng SimpliAmp Thermocycler (Hệ thống sinh học ứng dụng, Mỹ). cADN được kết hợp với nước không chứa nucleaza (Thermo Fisher Scientific Inc), mẫu dò TaqMan (thử nghiệm biểu hiện gen TaqMan (FAM), Thermo Fisher Scientific Inc) và hỗn hợp chính biểu hiện gen TaqMan (Thermo Fisher Scientific Inc) trong vi đĩa 384 giếng (Tấm phản ứng quang học 384 giếng MicroAmp, Thermo Fisher Scientific Inc). qPCR được thực hiện trên hệ thống PCR thời gian thực ViiA 7 (Hệ thống sinh học ứng dụng, Mỹ). Các thiết lập là 1: 2phút, 50°C; 2: 10phút, 95°C; 3: 15giây, 95°C; 4: 1 phút, 60°C. Bước 3 và 4 được lặp lại trong 45 chu kỳ. Phân tích dữ liệu được thực hiện sử dụng phần mềm ViiA 7, phần mềm phân tích dữ liệu qPCR được thực hiện bằng phần mềm MS Excel và GraphPad Prism.

Sự hấp thụ chất cản quang của gan được đo bằng hình ảnh cộng hưởng từ (MRI)

Các phương pháp thực hiện MRI đã được áp dụng từ công bố của Egger *et al* (Egger *et al*., (2015) J Magn Reson Imaging, 41: 829-840). Các thí nghiệm được thực hiện trên hệ thống MRI 7-T Bruker Biospec (Bruker Biospin, Ettlingen, Đức). Trong quá trình thu nhận tín hiệu MRI, chuột được đặt ở tư thế nằm ngửa trong giá đỡ Plexiglas. Nhiệt độ cơ thể được giữ ở 37 ± 1°C sử dụng đệm sưởi. Sau một thời gian ngắn khởi đầu, gây mê được duy trì với xấp xỉ 1,4% isofluran trong hỗn hợp O_2/N_2O (1:2), được dùng qua nón mũi. Tất cả các phép đo được thực hiện trên động vật thở tự nhiên; cả kích hoạt tim và hô hấp đều không được áp dụng.

Sau khi đặt chuột vào máy quét, các hình ảnh nhanh do thám được thu thập cho các mục đích cục bộ hóa. Các phân tích tưới máu được thực hiện sử dụng tác nhân nội mạch chứa các hạt nano sắt oxit siêu thuận từ (SPIO) (Endorem®, Guerbet, Pháp). Endorem® được tiêm trong tĩnh mạch ở tiêm bolus trong 1,2 giây vào con vật bị AKI (ở 24 giờ sau khi cảm ứng bệnh) hoặc sau sự hoạt động Giả bộ (các con vật sau thủ thuật cắt bỏ thận 24 giờ). Một liều bolus đầu tiên được sử dụng trong 1,2 giây, kết hợp với việc thu nhận tuần tự các hình ảnh cộng hưởng phẳng ở độ phân giải 400 mili giây/hình

ảnh. Sau khi thu được 25 hình ảnh đường cơ sở, liều bolus thứ hai được tiêm trong 1,2 giây và 575 hình ảnh khác được thu thập sau khi tiêm bolus, tạo thành tổng cộng 600 hình ảnh thu được trong 4 phút. Chất tương phản siêu thuận từ gây ra những thay đổi cục bộ về tính nhạy cảm, dẫn đến sự suy giảm tín hiệu tỷ lệ với sự tưới máu của thận. Đối với một loạt hình ảnh, cường độ tín hiệu được đánh giá trên các vùng quan tâm (các ROI) nằm ở vỏ não/dải ngoài của tủy ngoài. Vị trí, hình dạng và kích thước của các ROI đã được lựa chọn cẩn thận để đảm bảo rằng chúng bao phủ gần như cùng một vùng, mặc dù các chuyển động của thận có thể gây ra do hô hấp. Các cường độ tín hiệu trung bình cho các hình ảnh trước khi tiêm cung cấp các cường độ cơ bản (S (0)). Các chỉ số tưới máu được xác định từ giá trị trung bình của các tỷ lệ sau (Rosen et al., (1990) Magn Reson Med., 14: 249-265):

$$-\ln[S(t)/S(0)] \sim TE.V.cT(t)$$

trong đó TE là thời gian dội âm, V là thể tích máu, và cT là nồng độ của chất cản quang.

Các hạt nano SPIO được sử dụng trong nghiên cứu có đường kính trung bình khoảng 150 nm và được các tế bào Kupffer trong gan tiếp nhận. Do đó, ngoài việc tưới máu cho thận, MRI cũng cho phép theo dõi sự hấp thu của các hạt nano trong gan, bằng cách phát hiện sự thay đổi độ tương phản được đánh giá trong các ROI được đặt trong gan.

9.4 Kết quả

Như được thể hiện trên Hình 15, các protein dung hợp FP330 (EGF-HSA-C1-C2; SEQ ID No: 42), FP278 (EGF-HSA-C1-C2-thẻ His; SEQ ID No: 44) và FP776 (EGF-HSA-C1-C2; SEQ ID No: 48) bảo vệ chức năng thận trong mô hình chấn thương thận cấp (AKI) này khi dùng một trong hai phương pháp i.p. (FP278) hoặc i.v. (FP330 và FP776). Sự bảo vệ này được phản ánh bởi sự phong bế việc tăng creatinin huyết thanh (sCr). Hình 15A cho thấy protein dung hợp FP278 ở cả hai liều thử nghiệm đã làm giảm đáng kể các mức creatinin huyết thanh ($p < 0,0001$) so với các động vật được điều trị bằng phương tiện và hiệu quả như MFG-E8 ở chó. Như thể hiện trong Hình 15B, protein dung hợp FP330 bảo vệ chức năng thận theo cách phụ thuộc vào liều và tương tự như vậy đối với protein dung hợp FP776 (Hình 15C), trong đó mức creatinin huyết thanh cũng bị chặn theo cách phụ thuộc liều.

Chức năng thận bị suy giảm cũng được phản ánh qua các mức nitơ urê trong máu (BUN) ở chuột được thử nghiệm và tác động của protein dung hợp FP278 đối với các mức BUN được thể hiện trong Hình 16.

Tóm lại, như được thể hiện trong các Hình 15 và 16, các protein dung hợp FP278, FP330 và FP776 được bảo vệ mạnh mẽ chống lại sự gia tăng các chỉ thị này được sử dụng để chẩn đoán lâm sàng suy thận. Hiệu quả quan sát được đã được xác nhận bởi mô học (không được hiển thị).

Hơn nữa, như thể hiện trong Hình 17 liều đơn của protein dung hợp FP278 bảo vệ các cơ quan ở xa khỏi đáp ứng giai đoạn cấp tính do AKI gây ra. AKI tạo ra rất nhiều đáp ứng mARN có thể đo được bằng qPCR trong các màng lọc của các cơ quan được tưới máu cao ở xa như lá lách, tim gan phổi và não. Các mARN điển hình gây ra tổn thương chọn lọc (NGAL, KIM-1), cảm ứng chemokin (không được thể hiện) hoặc cảm ứng protein đáp ứng giai đoạn cấp tính như amyloid A (SAA) trong huyết thanh. Hình 17A và 17B minh họa ví dụ sự đáp ứng gây ra bởi AKI như vậy (dạng tính bột A huyết thanh (SAA)) ở tim và phổi chuột mà được phong bế một cách tiềm năng và được trở về mức Giả bộ sau khi một lần tiêm đơn lẻ protein dung hợp.

Gan hấp thu chất cản quang Endorem® theo thời gian được thể hiện trong Hình 18. Các con vật bị AKI thể hiện sự hấp thụ giảm đáng kể chất cản quang bởi gan (đích = tế bào Kupffer) so với các con vật Giả bộ. Việc điều trị FP776 (liều dự phòng 1,26 mg/kg, -30 phút trước khi khởi phát AKI, hoặc liều trị liệu 2 mg/kg, + 5 giờ sau khi khởi phát chấn thương tái tưới máu do thiếu máu cục bộ) được bảo vệ khỏi sự hao hụt tích tụ chất cản quang trong gan của chuột AKI . Những kết quả này cho thấy rằng trong mô hình chuột này, AKI gây ra sự suy giảm đáng kể khả năng thanh thải hạt qua trung gian tế bào Kupffer nội sinh và AKI gây ra rối loạn vi mạch ảnh hưởng đến sự tích tụ của chất cản quang hạt sắt trong gan. Việc điều trị bằng protein dung hợp FP776 được bảo vệ khỏi sự hao hụt độ thanh thải và khỏi rối loạn vi mạch, và thậm chí còn tăng cường sự hấp thu chất cản quang ở cả hai liều được thử nghiệm, khi so sánh với động vật giả.

Các protein dung hợp trị liệu, ví dụ theo Phương án 19 (ví dụ SEQ ID NO: 80) hoặc Phương án 20 (ví dụ SEQ ID NO: 82), thúc đẩy sự bám dính tế bào integrin av và thúc đẩy quá trình tiêu hủy tế bào chết giống với FPJ776 khi được thử nghiệm ở các thí nghiệm nêu trên. Do đó, chúng thích hợp cho các sử dụng trị liệu được bộc lộ ở đây.

Kết hợp lại với nhau, các dữ liệu này chứng minh rằng các protein dung hợp theo sáng chế, ví dụ, với việc chèn miền HSA, có chức năng và hiệu quả và do đó có thể được sử dụng làm phương pháp trị liệu.

Cần hiểu rằng các ví dụ và các phương án được mô tả ở đây chỉ nhằm mục đích minh họa và các sửa đổi hoặc cải biến khác theo đó sẽ được đề xuất cho các chuyên gia trong lĩnh vực và được đưa vào trong tinh thần và nội dung của sáng chế và phạm vi của các yêu cầu bảo hộ kèm theo. Tất cả các tài liệu công bố, bằng sáng chế, và đơn sáng chế được viện dẫn ở đây chỉ được kết hợp để tham khảo vì bất kỳ mục đích nào.

Danh Mục Trình Tự

Đơn sáng chế này có chứa Danh Mục Trình Tự mà được nộp theo hình thức điện tử dưới định dạng ASCII và được kết hợp toàn bộ trong bản mô tả này chỉ để tham khảo. Bản sao ASCII này, được tạo ngày 31 tháng 08 năm 2020, có tên là PAT058332_SL.txt và có kích thước 653193 byte.

YÊU CẦU BẢO HỘ

1. Protein dung hợp trị liệu để tăng cường quá trình tiêu hủy tế bào chết bao gồm miền liên kết integrin, miền liên kết phosphatidylserin (PS) và miền hòa tan, trong đó miền hòa tan được chèn ở giữa miền liên kết integrin và miền liên kết PS; và trong đó miền liên kết integrin liên kết với integrin, và trong đó miền hòa tan là HSA và có trình tự axit amin của SEQ ID NO: 4, hoặc tương đồng trình tự ít nhất 90% với nó.

2. Protein dung hợp theo điểm 1, trong đó protein dung hợp này có trình tự axit amin của SEQ ID NO: 42, hoặc tương đồng trình tự ít nhất 90% với nó.

3. Protein dung hợp theo điểm bất kỳ trong số các điểm nêu trên, trong đó protein dung hợp này có trình tự axit amin của SEQ ID NO: 44, hoặc tương đồng trình tự ít nhất 90% với nó.

4. Protein dung hợp theo điểm bất kỳ trong số các điểm nêu trên, trong đó protein dung hợp này có trình tự axit amin của SEQ ID NO: 47, hoặc tương đồng trình tự ít nhất 90% với nó.

5. Protein dung hợp theo điểm bất kỳ trong số các điểm nêu trên, trong đó protein dung hợp này có trình tự axit amin của SEQ ID NO: 48, hoặc tương đồng trình tự ít nhất 90% với nó.

6. Protein dung hợp theo điểm bất kỳ trong số các điểm nêu trên, trong đó protein dung hợp này có trình tự axit amin của SEQ ID NO: 80, hoặc tương đồng trình tự ít nhất 90% với nó.

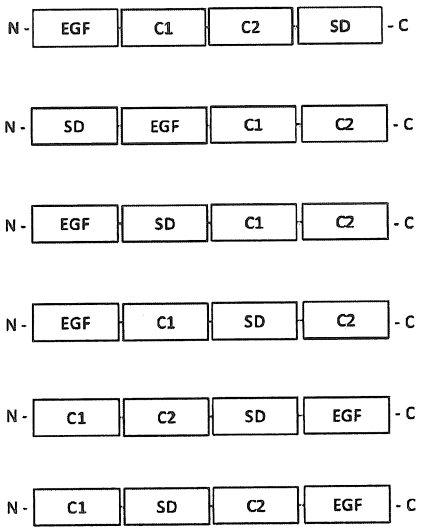
7. Protein dung hợp theo điểm bất kỳ trong số các điểm nêu trên, trong đó protein dung hợp này có trình tự axit amin của SEQ ID NO: 82, hoặc tương đồng trình tự ít nhất 90% với nó.

8. Axit nucleic được phân lập mã hóa trình tự axit amin theo điểm bất kỳ trong số các điểm từ 2 đến 7.

9. Vectơ tách dòng hoặc biểu hiện chứa axit nucleic theo điểm 8.

10. Tế bào chủ tái tổ hợp thích hợp để sản xuất protein dung hợp trị liệu, chứa một hoặc nhiều vectơ tách dòng hoặc biểu hiện theo điểm 9 và tùy ý, các tín hiệu tiết.

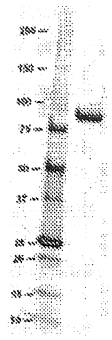
HÌNH 1


N - EGF C1 C2 SD - C
N - SD EGF C1 C2 - C
N - EGF SD C1 C2 - C
N - EGF C1 SD C2 - C
N - C1 C2 SD EGF - C
N - C1 SD C2 EGF - C

HÌNH 2

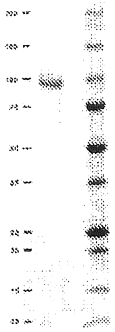
HÌNH 2A

EGF-HSA-C1-C2
(FP330)



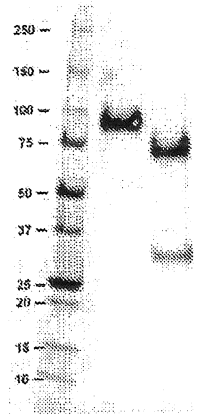
HÌNH 2B

EGF-HSA-C1-C2 của EDIL3
(FP050)



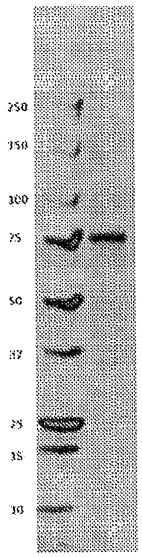
HÌNH 2C

EGF-Fc(KiH)-C1-C2
không giảm và giảm
(FP071 + Fc-IgG1 (lỗ))

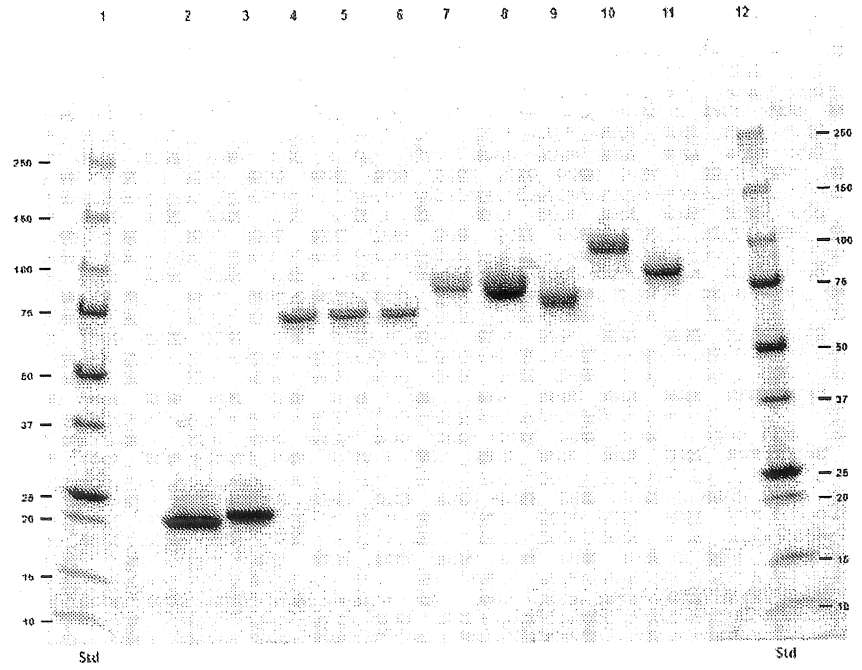


HÌNH 2D

EGF-HSA-C1
(FP260)



HÌNH 2E: Phân tích SDS-PAGE của các thể dung hợp MFG-E8, HSA bị cắt cụt và các protein khảm MFG-E8/EDIL-3

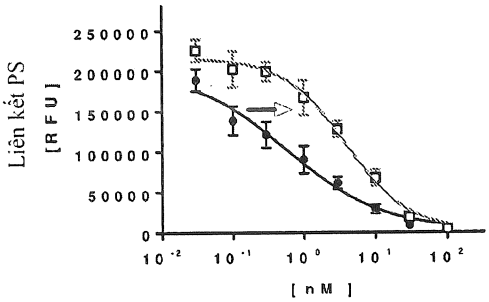


Đường 1, 12:	Chỉ thị khối lượng phân tử (Biorad, protein Precision Plus)		
Đường 2:	His6_EGF[MFG-E8]_C1[MFG-E8]		23,87kDa
Đường 3:	EGF[MFG-E8]_C1[MFG-E8]_His6	SEQ ID 115	23,87kDa
Đường 4:	EGF[MFG-E8]_HSA_C1[MFG-E8]	SEQ ID 117	90,38kDa
Đường 5:	EGF[MFG-E8]_HSA_C1[MFG-E8]	SEQ ID 74	89,27kDa
Đường 6:	EGF[MFG-E8]_HSA_C1[MFG-E8]	SEQ ID 73	88,72kDa
Đường 7:	EGF[EDIL3]_HSA_C1[EDIL3]	SEQ ID 71	98,22kDa
Đường 8:	EGF[EDIL3]_HSA_C2[EDIL3]	SEQ ID 135	98,20kDa
Đường 9:	EGF[MFG-E8]_HSA_C2[MFG-E8]	SEQ ID 137	88,45kDa
Đường 10:	EGF[EDIL3]_HSA_C1_C2[MFG-E8]	SEQ ID 80	115,67kDa
Đường 11:	EGF[MFG-E8]_HSA_C1_C2[EDIL3]	SEQ ID 82	107,32kDa

HÌNH 3

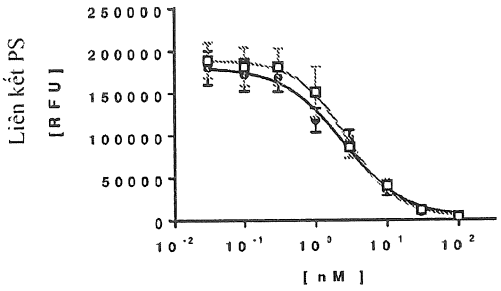
HÌNH 3A

WT MFG-E8 (EGF-C1-C2)



HÌNH 3B

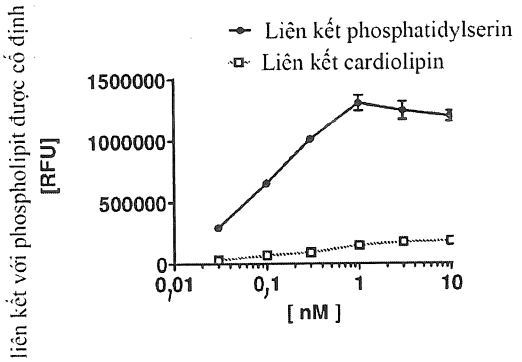
FP278 (EGF-HSA-C1-C2-thẻ His)



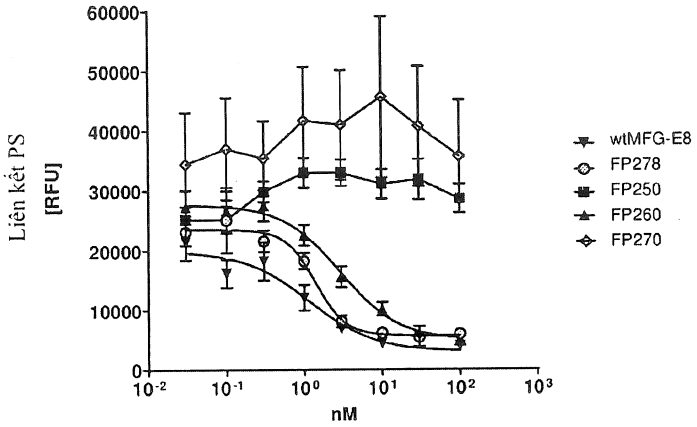
Sự pha loãng hoàn thành trong — Các đĩa không liên kết — Các đĩa polypropylen

HÌNH 4

HÌNH 4A

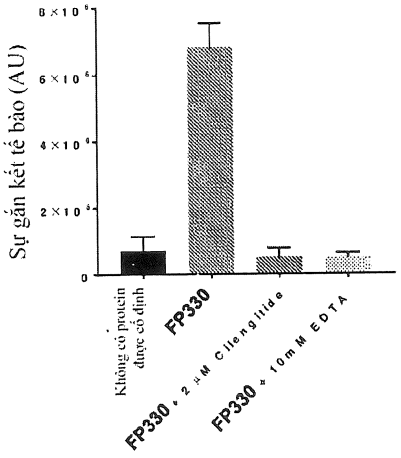


HÌNH 4B

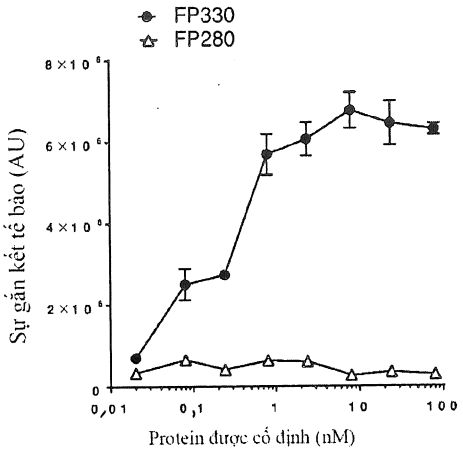


HÌNH 5

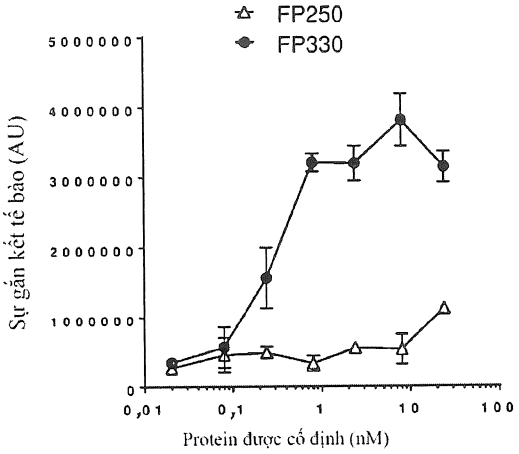
HÌNH 5A



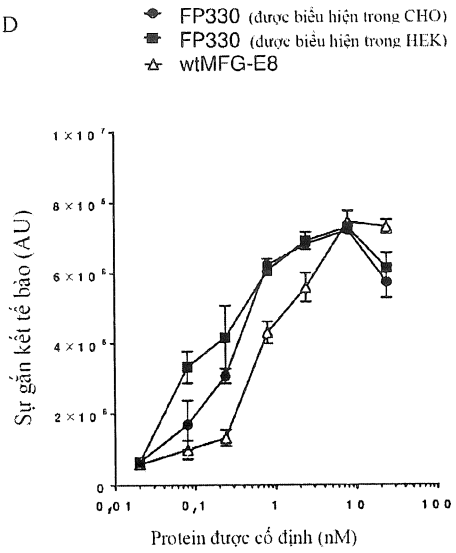
HÌNH 5B



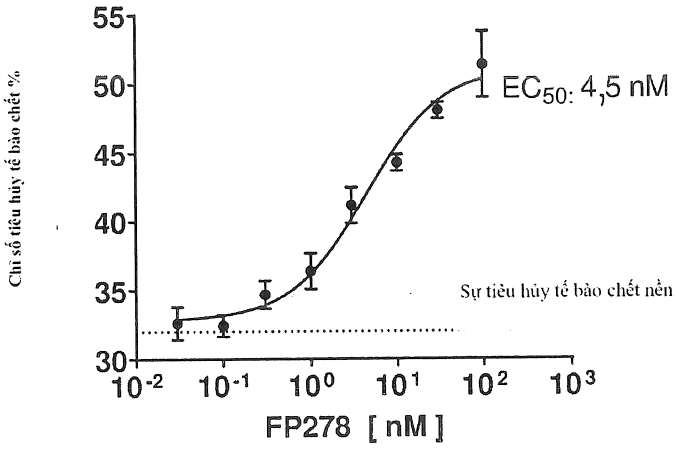
HÌNH 5C



HÌNH 5D

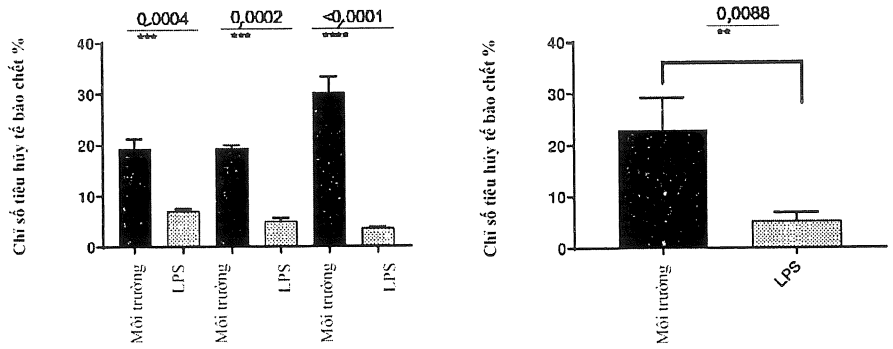


HÌNH 6

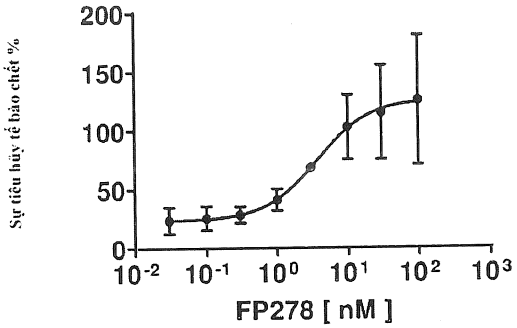

Chỉ số tiêu hủy tế bào chết %
55
50
45
40
35
30
EC50: 4,5 nM
Sự tiêu hủy tế bào chết nền
10-2
10-1
100
101
102
103
FP278 [nM]

HÌNH 7

HÌNH 7A

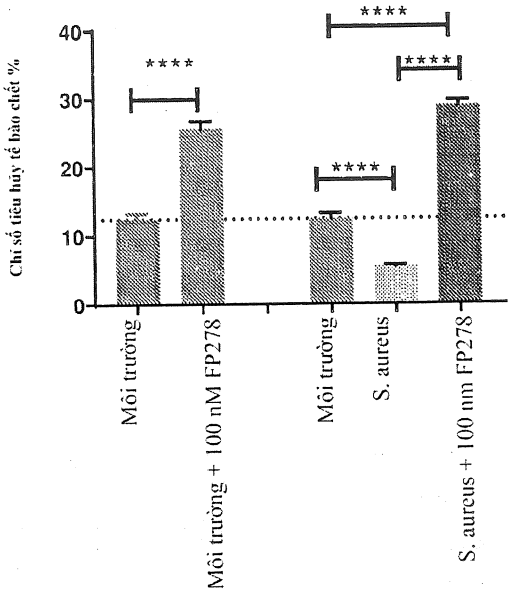


HÌNH 7B

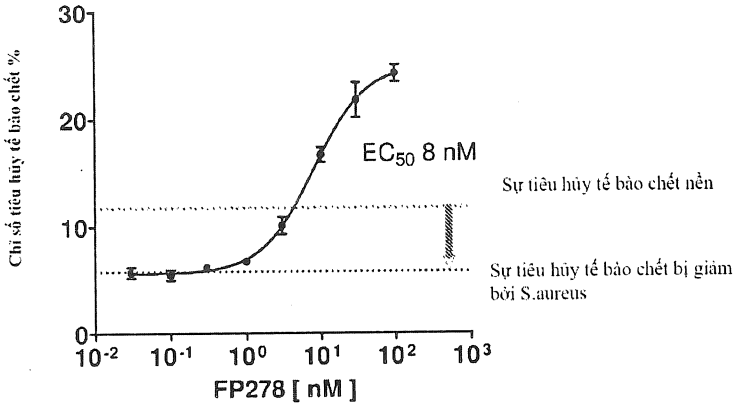


HÌNH 8

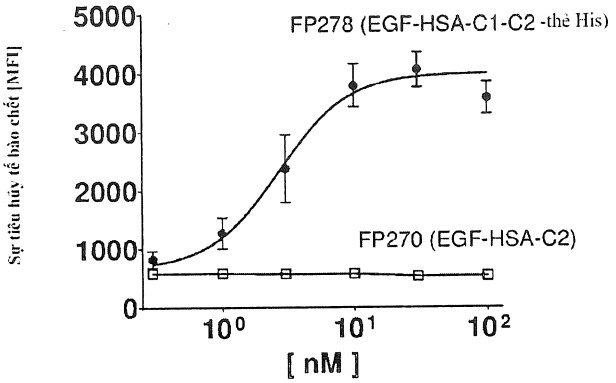
HÌNH 8A



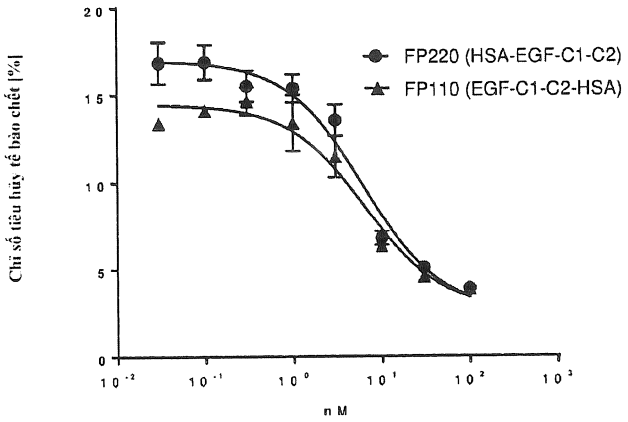
HÌNH 8B



HÌNH 9

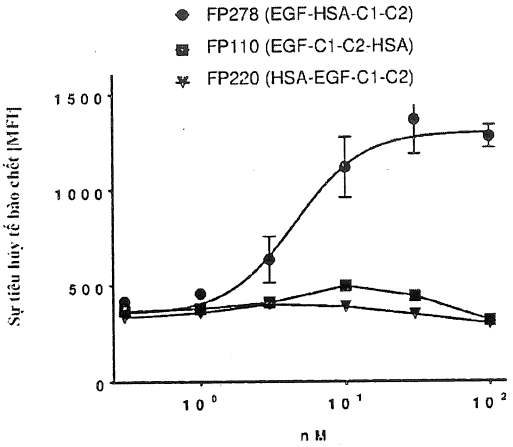


HÌNH 10

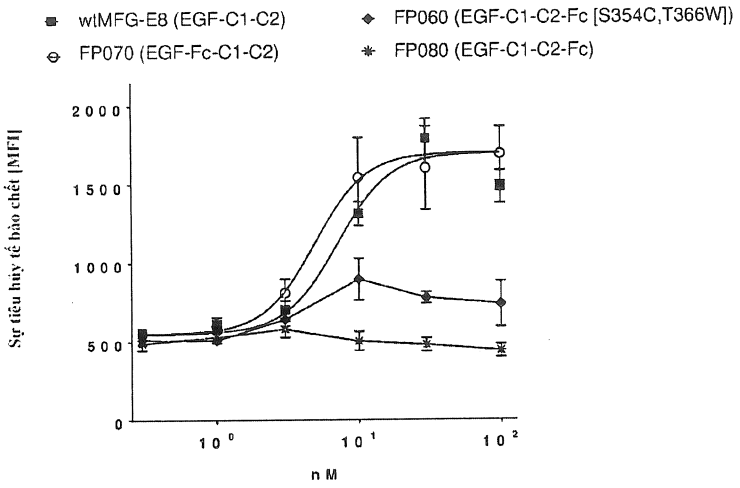


HÌNH 11

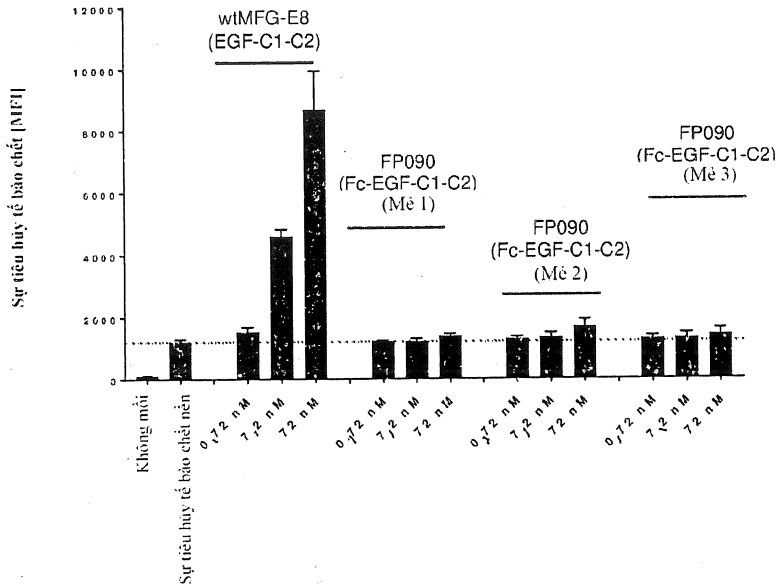
HÌNH 11A



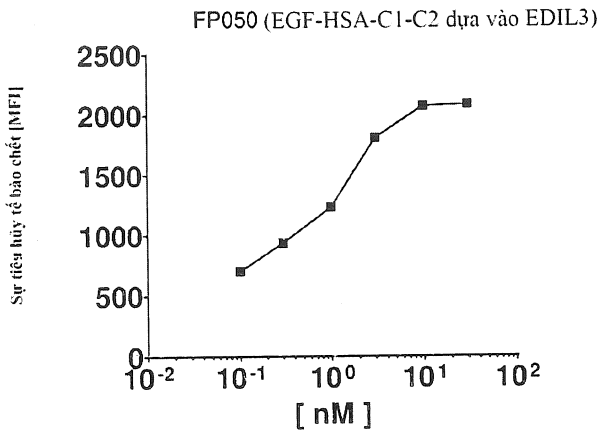
HÌNH 11B



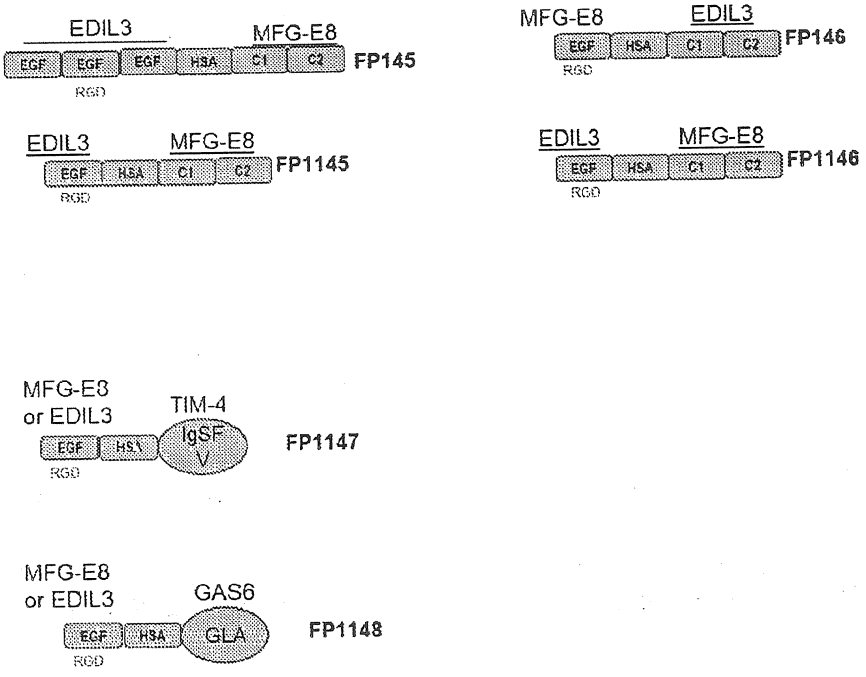
HÌNH 11C



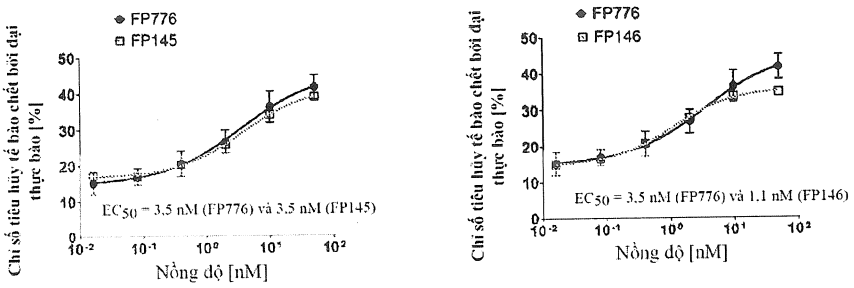
HÌNH 11D


FP050 (EGF-HSA-C1-C2 dựa vào EDIL3)
Sự tiêu hủy tế bào chết [MFI]
2500
2000
1500
1000
500
0
10-2
10-1
100
101
102
[nM]

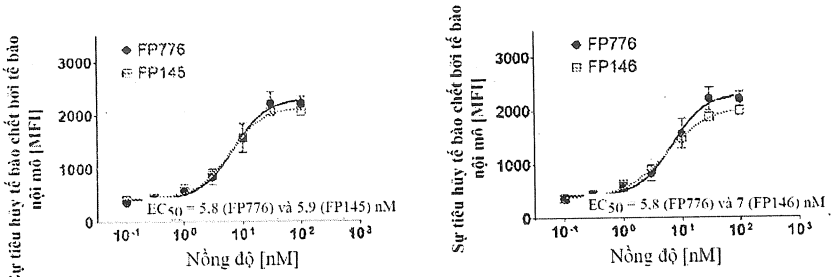
HÌNH 11E


EDIL3
MFG-E8
EGF
EGF
EGF
HSA
C1
C2
RGD
FP145
MFG-E8
EDIL3
EGF
HSA
C1
C2
RGD
FP146
EDIL3
MFG-E8
EGF
HSA
C1
C2
RGD
FP1145
EDIL3
MFG-E8
EGF
HSA
C1
C2
RGD
FP1146
MFG-E8
or EDIL3
TIM-4
EGF
HSA
IgSF
V
RGD
FP1147
MFG-E8
or EDIL3
GAS6
EGF
HSA
GLA
RGD
FP1148

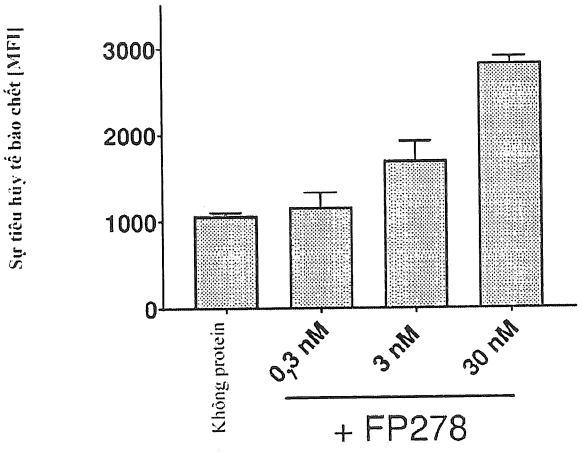
HÌNH 11F



HÌNH 11G

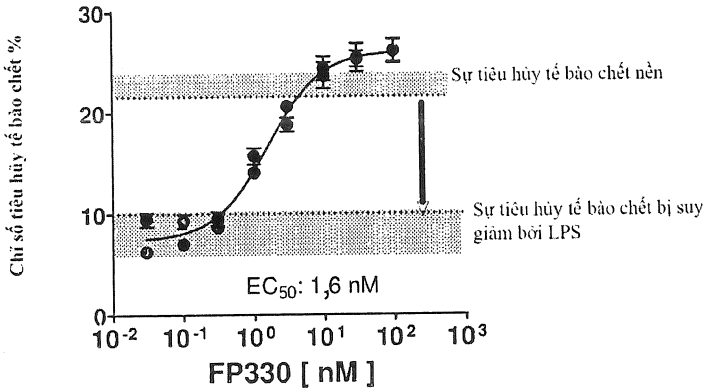


HÌNH 12

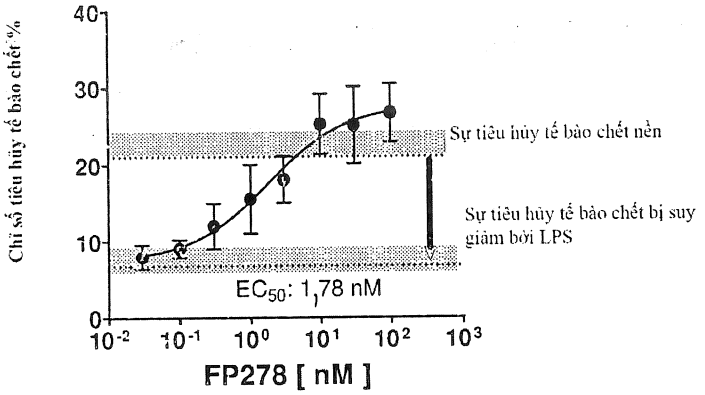

Sự tiêu hủy tế bào chết [MFI]
3000
2000
1000
0
Không protein
0,3 nM
3 nM
30 nM
+ FP278

HÌNH 13

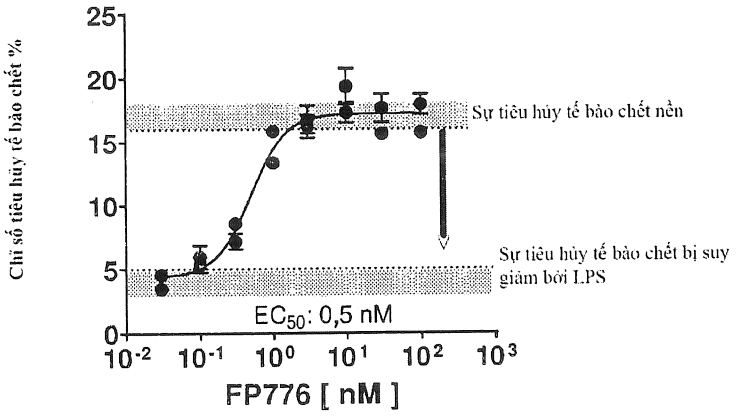
HÌNH 13A



HÌNH 13B

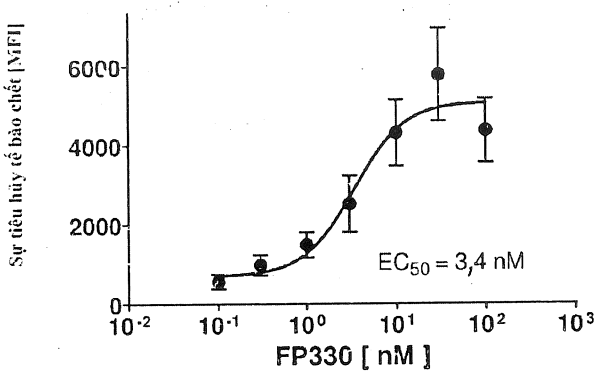


HÌNH 13C

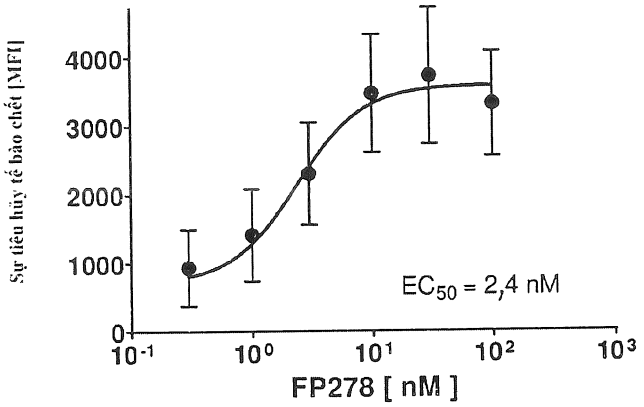


HÌNH 14

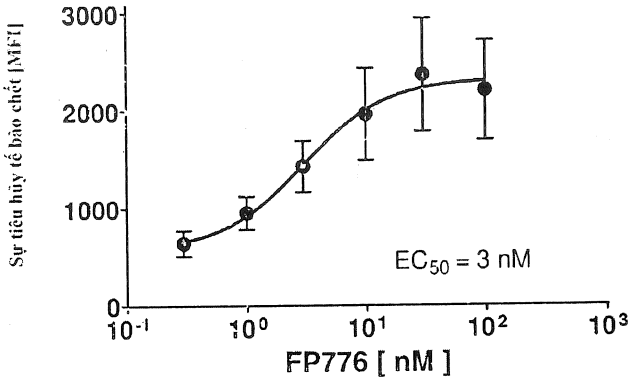
HÌNH 14A



HÌNH 14B

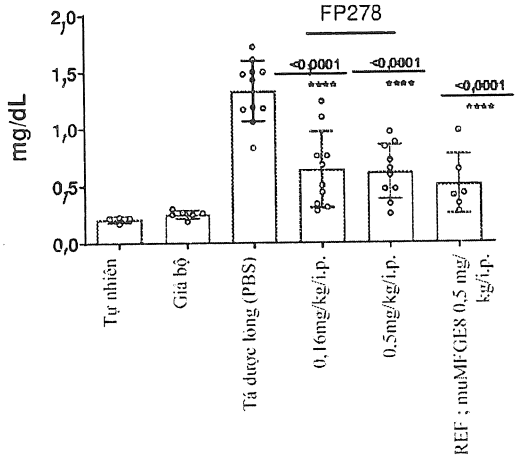


HÌNH 14C

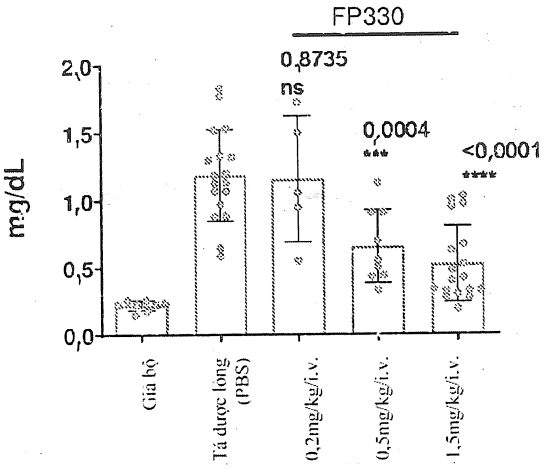


HÌNH 15

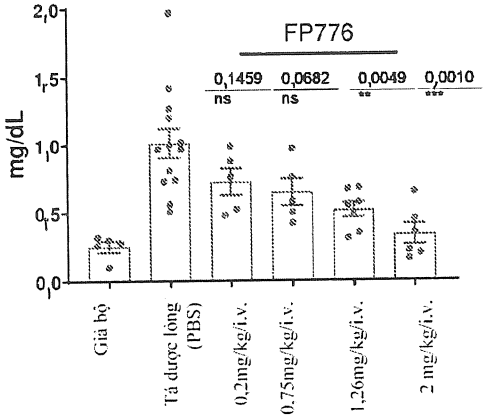
HÌNH 15A



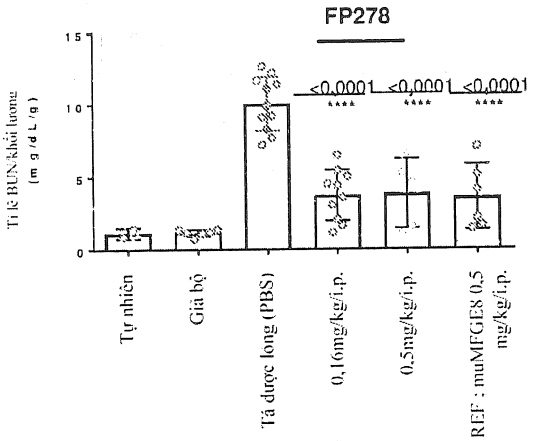
HÌNH 15B



HÌNH 15C



HÌNH 16

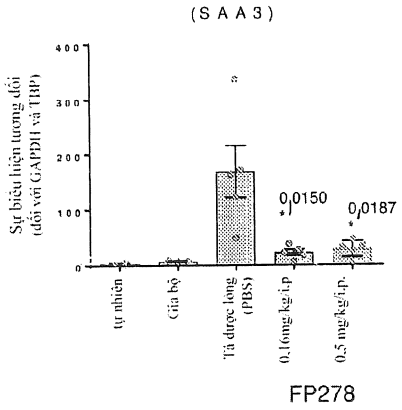


HÌNH 17

HÌNH 17A

Tim

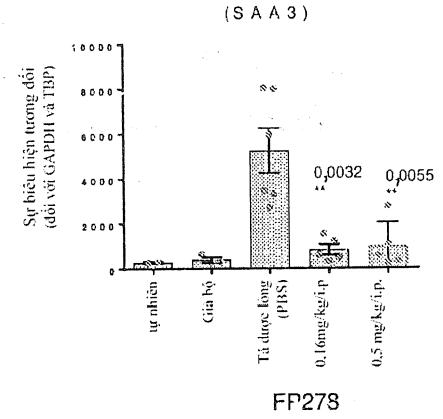
Amyloid A huyết thanh được cảm ứng bởi AKI



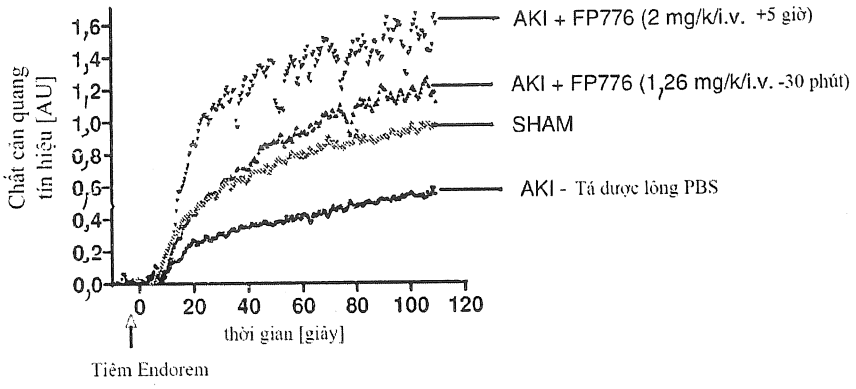
HÌNH 17B

Phổi

Amyloid A huyết thanh được cảm ứng bởi AKI



HÌNH 18


Chất cản quang
tín hiệu [AU]
1,6
1,4
1,2
1,0
0,8
0,6
0,4
0,2
0,0
0
20
40
60
80
100
120
thời gian [giây]
Tiêm Endorem
AKI + FP776 (2 mg/k/i.v. +5 giờ)
AKI + FP776 (1,26 mg/k/i.v. -30 phút)
SHAM
AKI - Tá dược lông PBS

DANH MỤC TRÌNH TỰ

<110> NOVARTIS AG

<120> PROTEIN DUNG HỢP TRỊ LIỆU ĐỂ TĂNG CƯỜNG QUÁ TRÌNH TIÊU HỦY TẾ BÀO CHẾT VÀ TẾ BÀO CHỦ TÁI TỔ HỢP THÍCH HỢP ĐỂ SẢN XUẤT PROTEIN NÀY

<130> PAT058332

<140>

<141>

<150> EP 19196045.9

<151> 2019-09-06

<160> 148

<170> PatentIn phiên bản 3.5

<210> 1

<211> 387

<212> PRT

<213> Homo sapiens

<400> 1

Met Pro Arg Pro Arg Leu Leu Ala Ala Leu Cys Gly Ala Leu Leu Cys
1 5 10 15

Ala Pro Ser Leu Leu Val Ala Leu Asp Ile Cys Ser Lys Asn Pro Cys
20 25 30

His Asn Gly Gly Leu Cys Glu Glu Ile Ser Gln Glu Val Arg Gly Asp
35 40 45

Val Phe Pro Ser Tyr Thr Cys Thr Cys Leu Lys Gly Tyr Ala Gly Asn
50 55 60

His Cys Glu Thr Lys Cys Val Glu Pro Leu Gly Leu Glu Asn Gly Asn
65 70 75 80

Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu
85 90 95

Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly
100 105 110

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

355 360 365

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

370 375 380

Leu Gly Cys

385

<210> 2

<211> 46

<212> PRT

<213> Homo sapiens

<400> 2

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu

1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys

20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys

35 40 45

<210> 3

<211> 318

<212> PRT

<213> Homo sapiens

<400> 3

Cys Val Glu Pro Leu Gly Leu Glu Asn Gly Asn Ile Ala Asn Ser Gln

1 5 10 15

Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp

20 25 30

Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp

35 40 45

Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu

50 55 60

Arg Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu

65 70 75 80

Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn
85 90 95
Gly His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu
100 105 110
Phe Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu
115 120 125
Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His
130 135 140
Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly
145 150 155 160
Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln
165 170 175
Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser
180 185 190
Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala
195 200 205
Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu
210 215 220
Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn
225 230 235 240
Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn
245 250 255
Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser
260 265 270
Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu
275 280 285
Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala
290 295 300
Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys
305 310 315

<210> 4

<211> 583

<212> PRT

<213> Homo sapiens

<400> 4

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1 5 10 15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
20 25 30

Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
35 40 45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
50 55 60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65 70 75 80

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
85 90 95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
100 105 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
115 120 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
130 135 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145 150 155 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
165 170 175

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
180 185 190

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
195 200 205

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro

210 215 220

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

225 230 235 240

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp

245 250 255

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser

260 265 270

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His

275 280 285

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

290 295 300

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

305 310 315 320

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg

325 330 335

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr

340 345 350

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu

355 360 365

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

370 375 380

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu

385 390 395 400

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro

405 410 415

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys

420 425 430

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys

435 440 445

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His

450 455 460

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
465 470 475 480
Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
485 490 495
Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
500 505 510
Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
515 520 525
Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
530 535 540
Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
545 550 555 560
Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
565 570 575
Ala Ala Ser Gln Ala Ala Leu
580

<210> 5
<211> 583
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 5
Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1 5 10 15
Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
20 25 30
Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
35 40 45
Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys

50 55 60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu

65 70 75 80

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro

85 90 95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu

100 105 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His

115 120 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg

130 135 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg

145 150 155 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala

165 170 175

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser

180 185 190

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu

195 200 205

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro

210 215 220

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

225 230 235 240

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp

245 250 255

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser

260 265 270

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His

275 280 285

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

290 295 300

Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
305 310 315 320
Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
325 330 335
Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
340 345 350
Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
355 360 365
Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
370 375 380
Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
385 390 395 400
Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
405 410 415
Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
420 425 430
Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
435 440 445
Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
450 455 460
Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
465 470 475 480
Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
485 490 495
Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
500 505 510
Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
515 520 525
Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
530 535 540
Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys

545 550 555 560

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val

565 570 575

Ala Ala Ser Gln Ala Ala Leu

580

<210> 6

<211> 206

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 6

Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe

1 5 10 15

Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr

20 25 30

Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser

35 40 45

Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala

50 55 60

Lys Arg Met Pro Cys Ala Glu Asp Cys Leu Ser Val Phe Leu Asn Gln

65 70 75 80

Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys

85 90 95

Cys Cys Thr Glu Ser Leu Val Asn Gly Arg Pro Cys Phe Ser Ala Leu

100 105 110

Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe

115 120 125

Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile

130 135 140

Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala
145 150 155 160
Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val
165 170 175
Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu
180 185 190
Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
195 200 205

<210> 7
<211> 217
<212> PRT
<213> Homo sapiens
<400> 7

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

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn

145 150 155 160

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu

165 170 175

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val

180 185 190

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln

195 200 205

Lys Ser Leu Ser Leu Ser Pro Gly Lys

210 215

<210> 8

<211> 217

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 8

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys

1 5 10 15

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val

20 25 30

Val Val Ala Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr

35 40 45

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu

50 55 60

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His

65 70 75 80

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys

85 90 95

Ala Leu Ala Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln

100 105 110

Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met

115 120 125

Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro

130 135 140

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn

145 150 155 160

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu

165 170 175

Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val

180 185 190

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln

195 200 205

Lys Ser Leu Ser Leu Ser Pro Gly Lys

210 215

<210> 9

<211> 217

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 9

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys

1 5 10 15

Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val

20 25 30

Val Val Ala Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr

35 40 45

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu

50 55 60

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His
65 70 75 80
Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys
85 90 95
Ala Leu Ala Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln
100 105 110
Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Cys Arg Glu Glu Met
115 120 125
Thr Lys Asn Gln Val Ser Leu Trp Cys Leu Val Lys Gly Phe Tyr Pro
130 135 140
Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn
145 150 155 160
Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu
165 170 175
Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val
180 185 190
Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln
195 200 205
Lys Ser Leu Ser Leu Ser Pro Gly Lys
210 215

<210> 10
<211> 217
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 10

Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys
1 5 10 15
Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val

20 25 30

Val Val Ala Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr

35 40 45

Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu

50 55 60

Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His

65 70 75 80

Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys

85 90 95

Ala Leu Ala Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln

100 105 110

Pro Arg Glu Pro Gln Val Cys Thr Leu Pro Pro Ser Arg Glu Glu Met

115 120 125

Thr Lys Asn Gln Val Ser Leu Ser Cys Ala Val Lys Gly Phe Tyr Pro

130 135 140

Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn

145 150 155 160

Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu

165 170 175

Val Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val

180 185 190

Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln

195 200 205

Lys Ser Leu Ser Leu Ser Pro Gly Lys

210 215

<210> 11

<211> 457

<212> PRT

<213> Homo sapiens

<400> 11

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro

1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr

20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu

35 40 45

Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly

50 55 60

Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr

65 70 75 80

Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn

85 90 95

Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr

100 105 110

Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly

115 120 125

Arg Asn Cys Gln Tyr Lys Cys Ser Gly Pro Leu Gly Ile Glu Gly Gly

130 135 140

Ile Ile Ser Asn Gln Gln Ile Thr Ala Ser Ser Thr His Arg Ala Leu

145 150 155 160

Phe Gly Leu Gln Lys Trp Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys Lys

165 170 175

Gly Leu Ile Asn Ala Trp Thr Ala Ala Glu Asn Asp Arg Trp Pro Trp

180 185 190

Ile Gln Ile Asn Leu Gln Arg Lys Met Arg Val Thr Gly Val Ile Thr

195 200 205

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

210 215 220

Ile Ala Tyr Ser Asn Asp Gly Lys Thr Trp Ala Met Tyr Lys Val Lys

225 230 235 240

Gly Thr Asn Glu Asp Met Val Phe Arg Gly Asn Ile Asp Asn Asn Thr

245 250 255

Pro Tyr Ala Asn Ser Phe Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg
260 265 270

Leu Tyr Pro Gln Val Cys Arg Arg His Cys Thr Leu Arg Met Glu Leu
275 280 285

Leu Gly Cys Glu Leu Ser Gly Cys Ser Glu Pro Leu Gly Met Lys Ser
290 295 300

Gly His Ile Gln Asp Tyr Gln Ile Thr Ala Ser Ser Ile Phe Arg Thr
305 310 315 320

Leu Asn Met Asp Met Phe Thr Trp Glu Pro Arg Lys Ala Arg Leu Asp
325 330 335

Lys Gln Gly Lys Val Asn Ala Trp Thr Ser Gly His Asn Asp Gln Ser
340 345 350

Gln Trp Leu Gln Val Asp Leu Leu Val Pro Thr Lys Val Thr Gly Ile
355 360 365

Ile Thr Gln Gly Ala Lys Asp Phe Gly His Val Gln Phe Val Gly Ser
370 375 380

Tyr Lys Leu Ala Tyr Ser Asn Asp Gly Glu His Trp Thr Val Tyr Gln
385 390 395 400

Asp Glu Lys Gln Arg Lys Asp Lys Val Phe Gln Gly Asn Phe Asp Asn
405 410 415

Asp Thr His Arg Lys Asn Val Ile Asp Pro Pro Ile Tyr Ala Arg His
420 425 430

Ile Arg Ile Leu Pro Trp Ser Trp Tyr Gly Arg Ile Thr Leu Arg Ser
435 440 445

Glu Leu Leu Gly Cys Thr Glu Glu Glu
450 455

<210> 12

<211> 1052

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 12

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45

Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly
50 55 60

Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr
65 70 75 80

Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn
85 90 95

Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr
100 105 110

Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly
115 120 125

Arg Asn Cys Gln Tyr Lys Gly Ser Asp Ala His Lys Ser Glu Val Ala
130 135 140

His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu
145 150 155 160

Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val
165 170 175

Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp
180 185 190

Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp
195 200 205

Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala
210 215 220

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

465 470 475 480

Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys

485 490 495

Cys Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu

500 505 510

Phe Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys

515 520 525

Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu

530 535 540

Val Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val

545 550 555 560

Glu Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His

565 570 575

Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val

580 585 590

Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg

595 600 605

Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe

610 615 620

Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala

625 630 635 640

Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu

645 650 655

Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys

660 665 670

Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala

675 680 685

Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe

690 695 700

Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly

705 710 715 720

Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Cys Ser Gly
725 730 735

Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn Gln Gln Ile Thr Ala
740 745 750

Ser Ser Thr His Arg Ala Leu Phe Gly Leu Gln Lys Trp Tyr Pro Tyr
755 760 765

Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp Thr Ala Ala
770 775 780

Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln Arg Lys Met
785 790 795 800

Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile Gly Ser Pro
805 810 815

Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp Gly Lys Thr
820 825 830

Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Asp Met Val Phe Arg
835 840 845

Gly Asn Ile Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe Thr Pro Pro
850 855 860

Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Val Cys Arg Arg His
865 870 875 880

Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser Gly Cys Ser
885 890 895

Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr Gln Ile Thr
900 905 910

Ala Ser Ser Ile Phe Arg Thr Leu Asn Met Asp Met Phe Thr Trp Glu
915 920 925

Pro Arg Lys Ala Arg Leu Asp Lys Gln Gly Lys Val Asn Ala Trp Thr
930 935 940

Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp Leu Leu Val
945 950 955 960

Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys Asp Phe Gly

965 970 975

His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser Asn Asp Gly

980 985 990

Glu His Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys Asp Lys Val

995 1000 1005

Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn Val Ile

1010 1015 1020

Asp Pro Pro Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp Ser

1025 1030 1035

Trp Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys

1040 1045 1050

<210> 13

<211> 3156

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 13

gacatctgcg accccaatcc ttgcgagaat ggcggcattt gtctgcctgg actggccgat 60

ggcagcttct cttgtgaatg ccccgatggc ttcacagacc ccaattgcag ctctgtggtg 120

gaagtggcca gcgacgagga agaacctaca agcgctggcc cctgcacacc caatccatgt 180

cataatggcg gaacctgcga gatcagcgag gcctacagag gcgatacctt catcggctac 240

gtgtgcaagt gccccagagg cttcaatggc atccactgcc agcacaacat caacgagtgc 300

gaggtggaac catgcaagaa cggcggcatc tgtaccgacc tggtggccaa ttactcttgc 360

gagtgccctg gcgagttcat gggcagaaac tgccagtaca agggatccga cgctcacaag 420

tctgaggtgg cccacagatt caaggacctg ggcgaagaga acttcaaggc cctggtgctg 480

atcgccttcg ctcagtatct gcagcagagc cctttcgagg accacgtgaa gctggtcaac 540

gaagtgaccg agttcgccaa gacctgtgtg gccgatgaga gcgccgagaa ctgtgacaag 600

agcctgcaca cactgttcgg cgacaagctg tgtaccgtgg ccacactgag agaaacctac 660

ggcgagatgg ccgactgctg tgccaagcaa gagcccgaga gaaacgagtg cttcctgcag 720

cacaaggacg acaaccccaa cctgcctaga ctcgtgcgac ccgaagtgga tgtgatgtgc 780

accgcctttc acgacaacga ggaaaccttc ctgaagaagt acctgtacga gatcgccaga 840

cggcacccct acttttatgc ccctgagctg ctgttcttcg ccaagcggta taaggccgcc 900

ttcaccgaat gttgccaggc cgctgataag gctgcctgtc tgctgcctaa gctggacgag 960

ctgagagatg agggcaaagc cagctctgcc aagcagagac tgaagtgcgc cagcctgcag 1020

aagttcggcg agagagcttt taaggcctgg gccgttgcca gactgagcca gagatttcct 1080

aaggccgagt ttgccgaggt gtccaagctc gtgaccgatc tgacaaaggt gcacaccgag 1140

tgctgtcacg gcgatctgct ggaatgtgcc gacgatagag ccgacctggc caagtatatc 1200

tgcgagaacc aggacagcat cagcagcaag ctgaaagagt gctgcgagaa gcccctgctg 1260

gaaaagtctc actgtatcgc cgaggtcgag aacgacgaga tgcctgctga tctgcctagc 1320

ctggccgccg atttcgtgga aagcaaggat gtgtgcaaga actacgccga ggccaaagat 1380

gtgtttctgg gcatgtttct gtatgagtac gcccgcagac accccgacta ttctgtggtt 1440

ctgctgctgc ggctggccaa gacatacgag acaaccctgg aaaaatgctg cgccgctgcc 1500

gatcctcacg agtgttatgc caaggtgttc gacgagttca agccactggt ggaagaaccc 1560

cagaacctga tcaagcagaa ctgcgagctg ttcgagcagc tgggcgagta caagttccag 1620

aatgccctgc tcgtgcggta caccaagaaa gtgcctcagg tgtccacacc tacactggtt 1680

gaggtgtccc ggaatctggg caaagtgggc agcaagtgtt gcaagcaccc tgaggccaag 1740

agaatgcctt gcgccgagga ttacctgagc gtggtgctga atcagctgtg cgtgctgcac 1800

gagaaaaccc ctgtgtccga cagagtgacc aagtgctgta ccgagagcct cgtgaacaga 1860

aggccttgct ttagcgccct ggaagtggac gagacatacg tgcccaaaga gttcaacgcc 1920

gagacattca ccttccacgc cgatatctgc accctgtccg agaaagagcg gcagatcaag 1980

aagcagacag ccctggtcga gctggttaag cacaagccca aggccaccaa agaacagctg 2040

aaggccgtga tggacgactt cgccgccttt gtcgagaagt gctgcaaggc cgacgacaaa 2100

gagacatgct tcgccgaaga gggcaagaaa ctggtggctg cctctcaggc tgctctcgga 2160

cttggaggaa gcggaggatc tggcggttcc ggaggaagtt gttctggccc tcttggcatc 2220

gaaggcggca tcatcagcaa tcagcagatc accgccagca gcacccacag agcactgttt 2280

ggactgcaga aatggtatcc ctactacgcc cggctgaaca agaagggcct gattaacgcc 2340

tggacagccg ccgagaatga cagatggccc tggattcaga tcaacctgca gcggaagatg 2400

agagtgaccg gcgttatcac acagggcgcc aaaagaatcg gcagccccga gtacatcaag 2460

agctacaaga tcgcctacag caacgacggc aagacctggg ccatgtacaa agtgaagggc 2520

accaacgagg acatggtgtt ccggggcaac atcgacaaca acacccctta cgccaacagc 2580

ttcacccctc ctatcaaggc ccagtacgtg cggctgtacc ctcaagtgtg cagaaggcac 2640

tgtaccctga gaatggaact gctgggctgc gaactgtctg gctgttctga gccactgggc 2700

atgaagtccg gccacatcca ggattaccag atcacagcct ccagcatctt cagaaccctg 2760

aacatggata tgttcacctg ggagccccgg aaggccagac tggataagca gggaaaagtg 2820

aatgcctgga ccagcggcca caacgaccag tctcaatggc tgcaagtgga cctgctggtg 2880

cccaccaaag tgaccggaat cattactcag ggcgcaaagg acttcggcca cgtgcagttt 2940

gtgggctcct acaagctggc ctactccaac gatggcgagc actggacagt gtaccaggac 3000

gagaagcagc gcaaggataa ggtgttccag ggaaacttcg ataacgatac ccaccggaag 3060

aacgtgatcg accctccaat ctacgccaga cacatcagaa tcctgccttg gtcttggtac 3120

ggcagaatca ccctgagatc cgagctgctg ggatgc 3156

<210> 14

<211> 596

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 14

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Cys Val
35 40 45

Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile Ala
50 55 60

Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val Pro
65 70 75 80

Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr Pro
85 90 95

Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg Arg

100 105 110

Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala Ser

115 120 125

His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly His

130 135 140

Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe Val

145 150 155 160

Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr Pro

165 170 175

Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr Ala

180 185 190

Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys Ala

195 200 205

Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile Thr

210 215 220

Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp Asn

225 230 235 240

Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp Val

245 250 255

Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly Ser

260 265 270

Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly

275 280 285

Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser

290 295 300

Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys Ile

305 310 315 320

Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe Glu

325 330 335

Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp His

340 345 350

Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys Gly Gly Gly Gly
355 360 365

Thr Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu
370 375 380

Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu
385 390 395 400

Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser
405 410 415

His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu
420 425 430

Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr
435 440 445

Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn
450 455 460

Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro
465 470 475 480

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln
485 490 495

Val Tyr Thr Leu Pro Pro Cys Arg Glu Glu Met Thr Lys Asn Gln Val
500 505 510

Ser Leu Trp Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val
515 520 525

Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro
530 535 540

Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr
545 550 555 560

Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val
565 570 575

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
580 585 590

Ser Pro Gly Lys

595

<210> 15

<211> 1788

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 15

ctggacatct gcagcaagaa cccctgccac aacggcggcc tgtgcgaaga gatcagccag 60

gaagtgcggg gcgacgtgtt ccccagctac acctgtacct gcctgaaggg ctacgccggc 120

aaccactgcg agactaagtg cgtggaaccc ctgggcatgg aaaacggcaa tattgccaac 180

agccagatcg ccgccagctc cgtgcgcgtg acctttctgg gactgcagca ctgggtgccc 240

gagctggcca gactgaacag agccggcatg gtgaacgcct ggacccccag cagcaacgac 300

gacaaccctt ggatccaggt gaacctgctg cggcggatgt gggtgacagg cgtggtgaca 360

cagggcgcca gcagactggc cagccacgag tacctgaagg cctttaaggt ggcctacagc 420

ctgaacggcc acgagttcga cttcatccac gacgtgaaca agaaacacaa agaatttgtg 480

ggcaactgga acaagaacgc cgtgcacgtg aacctgttcg agacacccgt ggaagcccag 540

tacgtgcggc tgtaccccac cagctgccac accgcctgca ccctgagatt cgagctgctg 600

ggctgcgagc tgaacggctg cgccaacccc ctgggcctga agaacaacag catccccgac 660

aagcagatca ccgcctccag cagctacaag acctggggcc tgcacctgtt cagctggaac 720

cccagctacg cccggctgga caagcagggc aacttcaacg cctgggtggc cggcagctac 780

ggcaacgacc agtggctgca ggtggacctg ggcagcagca aagaagtgac cggcatcatc 840

acccaggggg ccagaaactt cggcagcgtg cagttcgtgg ccagctacaa agtggcctac 900

tccaacgaca gcgccaactg gaccgagtac caggaccccc ggaccggcag ctccaagatc 960

ttccccggca actgggacaa ccacagccac aagaagaatc tgttcgaaac ccccatcctg 1020

gccagatacg tgcggatcct gcccgtggcc tggcacaacc ggatcgccct gagactggaa 1080

ctgctgggat gtgggggagg cggtaccgac aagacccaca cctgcccccc ctgcccagcc 1140

ccagagctgc tgggcggacc ctccgtgttc ctgttccccc ccaagcccaa ggacaccctg 1200

atgatcagca ggacccccga ggtgacctgc gtggtggtgg acgtgagcca cgaggaccca 1260

gaggtgaagt tcaactggta cgtggacggc gtggaggtgc acaacgccaa gaccaagccc 1320

agagaggagc agtacaacag cacctacagg gtggtgtccg tgctgaccgt gctgcaccag 1380

gactggctga acggcaagga atacaagtgc aaggtctcca acaaggccct gccagccccc 1440

atcgaaaaga ccatcagcaa ggccaagggc cagccacggg agccccaggt gtacaccctg 1500

ccccccctgcc gggaggagat gaccaagaac caggtgtccc tgtggtgtct ggtgaagggc 1560

ttctacccca gcgacatcgc cgtggagtgg gagagcaacg gccagcccga gaacaactac 1620

aagaccaccc ccccagtgct ggacagcgac ggcagcttct tcctgtacag caagctgacc 1680

gtggacaagt ccaggtggca gcagggcaac gtgttcagct gcagcgtgat gcacgaggcc 1740

ctgcacaacc actacaccca gaagagcctg agcctgtccc ccggcaag 1788

<210> 16

<211> 613

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 16

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
50 55 60

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
65 70 75 80

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
85 90 95

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
100 105 110

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr

115 120 125

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly

130 135 140

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile

145 150 155 160

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

165 170 175

Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser

180 185 190

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu

195 200 205

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro

210 215 220

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val

225 230 235 240

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met

245 250 255

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser

260 265 270

Pro Gly Lys Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Cys

275 280 285

Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile

290 295 300

Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val

305 310 315 320

Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr

325 330 335

Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg

340 345 350

Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala

355 360 365

Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly
370 375 380
His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe
385 390 395 400
Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr
405 410 415
Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr
420 425 430
Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys
435 440 445
Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile
450 455 460
Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp
465 470 475 480
Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp
485 490 495
Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly
500 505 510
Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe
515 520 525
Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp
530 535 540
Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys
545 550 555 560
Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe
565 570 575
Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp
580 585 590
His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys Gly Ser His
595 600 605
His His His His His

610

<210> 17

<211> 1839

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 17

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccgacaag acccacacct gccccccctg cccagcccca 180

gagctgctgg gcggaccctc cgtgttcctg ttcccccca agcccaagga caccctgatg 240

atcagcagga cccccgaggt gacctgcgtg gtggtggacg tgagccacga ggacccagag 300

gtgaagttca actggtacgt ggacggcgtg gaggtgcaca acgccaagac caagcccaga 360

gaggagcagt acaacagcac ctacagggtg gtgtccgtgc tgaccgtgct gcaccaggac 420

tggctgaacg gcaaggaata caagtgcaag gtctccaaca aggccctgcc agcccccatc 480

gaaaagacca tcagcaaggc caagggccag ccacgggagc cccaggtgta caccctgccc 540

ccctcccggg aggagatgac caagaaccag gtgtccctga cctgtctggt gaagggcttc 600

taccccagcg acatcgccgt ggagtgggag agcaacggcc agcccgagaa caactacaag 660

accacccccc cagtgctgga cagcgacggc agcttcttcc tgtacagcaa gctgaccgtg 720

gacaagtcca ggtggcagca gggcaacgtg ttcagctgca gcgtgatgca cgaggccctg 780

cacaaccact acacccagaa gagcctgagc ctgtcccccg gcaagggagg aagcggagga 840

tctggcggtt ccggaggctc ttgtgtggaa cccctcggca tggaaaacgg caatatcgcc 900

aatagccaga ttgccgccag cagcgtcaga gtgacatttc tgggactgca gcactgggtg 960

cccgagctgg ctagactgaa tagagccggc atggtcaacg cctggacacc cagcagcaac 1020

gacgataacc cttggattca agtgaacctg ctgcggcgta tgtgggtcac aggtgttgtt 1080

acacagggcg cctctagact ggccagccac gagtatctga aggcctttaa ggtggcctac 1140

agcctgaacg gccacgagtt cgacttcatc cacgacgtga acaagaagca caaagagttt 1200

gtcggcaact ggaacaagaa cgccgtgcac gtgaacctgt tcgagacacc tgtggaagcc 1260

cagtacgtgc ggctgtaccc tacaagctgt cacaccgcct gcactctgag attcgaactg 1320

ctgggatgcg agctgaacgg ctgtgctaat cctctgggcc tgaagaacaa cagcatcccc 1380

gataagcaga tcaccgccag ctccagctat aagacatggg gcctgcacct gttcagctgg 1440

aacccttctt acgccagact ggacaagcag ggcaacttca atgcttgggt ggccggcagc 1500

tacggcaatg atcagtggct gcaagtggac ctgggcagca gcaaagaagt gacaggcatc 1560

atcacccagg gcgccagaaa tttcggcagc gtgcagtttg tggccagcta caaagtggcc 1620

tactccaacg acagcgccaa ctggaccgag tatcaggacc ctagaaccgg cagctccaag 1680

atcttccccg gcaattggga caaccacagc cacaagaaga atctgttcga aacccctatc 1740

ctggccagat atgtgcgcat tctgcccgtg gcctggcaca acagaattgc cctgagactg 1800

gaactgctcg gctgtggctc tcaccaccac catcaccat 1839

<210> 18

<211> 613

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 18

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
50 55 60

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
65 70 75 80

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
85 90 95

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
100 105 110

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

355 360 365

Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly

370 375 380

His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe

385 390 395 400

Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr

405 410 415

Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr

420 425 430

Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys

435 440 445

Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile

450 455 460

Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp

465 470 475 480

Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp

485 490 495

Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly

500 505 510

Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe

515 520 525

Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp

530 535 540

Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys

545 550 555 560

Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe

565 570 575

Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp

580 585 590

His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys Gly Ser His

595 600 605

His His His His His

610

<210> 19

<211> 1839

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 19

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccgacaag acccacacct gccccccctg cccagcccca 180

gagctgctgg gcggaccctc cgtgttcctg ttcccccccа agcccaagga caccctgatg 240

atcagcagga cccccgaggt gacctgcgtg gtggtggacg tgagccacga ggacccagag 300

gtgaagttca actggtacgt ggacggcgtg gaggtgcaca acgccaagac caagcccaga 360

gaggagcagt acaacagcac ctacagggtg gtgtccgtgc tgaccgtgct gcaccaggac 420

tggctgaacg gcaaggaata caagtgcaag gtctccaaca aggccctgcc agcccccatc 480

gaaaagacca tcagcaaggc caagggccag ccacgggagc cccaggtgta caccctgccc 540

ccctgccggg aggagatgac caagaaccag gtgtccctgt ggtgtctggt gaagggcttc 600

taccccagcg acatcgccgt ggagtgggag agcaacggcc agcccgagaa caactacaag 660

accacccccc cagtgctgga cagcgacggc agcttcttcc tgtacagcaa gctgaccgtg 720

gacaagtcca ggtggcagca gggcaacgtg ttcagctgca gcgtgatgca cgaggccctg 780

cacaaccact acacccagaa gagcctgagc ctgtcccccg gcaagggagg aagcggagga 840

tctggcggtt ccggaggctc ttgtgtggaa cccctcggca tggaaaacgg caatatcgcc 900

aatagccaga ttgccgccag cagcgtcaga gtgacatttc tgggactgca gcactgggtg 960

cccgagctgg ctagactgaa tagagccggc atggtcaacg cctggacacc cagcagcaac 1020

gacgataacc cttggattca agtgaacctg ctgcggcgta tgtgggtcac aggtgttgtt 1080

acacagggcg cctctagact ggccagccac gagtatctga aggcctttaa ggtggcctac 1140

agcctgaacg gccacgagtt cgacttcatc cacgacgtga acaagaagca caaagagttt 1200

gtcggcaact ggaacaagaa cgccgtgcac gtgaacctgt tcgagacacc tgtggaagcc 1260

cagtacgtgc ggctgtaccc tacaagctgt cacaccgcct gcactctgag attcgaactg 1320

ctgggatgcg agctgaacgg ctgtgctaat cctctgggcc tgaagaacaa cagcatcccc 1380

gataagcaga tcaccgccag ctccagctat aagacatggg gcctgcacct gttcagctgg 1440

aacccttctt acgccagact ggacaagcag ggcaacttca atgcttgggt ggccggcagc 1500

tacggcaatg atcagtggct gcaagtggac ctgggcagca gcaaagaagt gacaggcatc 1560

atcacccagg gcgccagaaa tttcggcagc gtgcagtttg tggccagcta caaagtggcc 1620

tactccaacg acagcgccaa ctggaccgag tatcaggacc ctagaaccgg cagctccaag 1680

atcttccccg gcaattggga caaccacagc cacaagaaga atctgttcga aacccctatc 1740

ctggccagat atgtgcgcat tctgcccgtg gcctggcaca acagaattgc cctgagactg 1800

gaactgctcg gctgtggctc tcaccaccac catcaccat 1839

<210> 20

<211> 613

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 20

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
50 55 60

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
65 70 75 80

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
85 90 95

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val

100 105 110

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr

115 120 125

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly

130 135 140

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile

145 150 155 160

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

165 170 175

Cys Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser

180 185 190

Leu Ser Cys Ala Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu

195 200 205

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro

210 215 220

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Val Ser Lys Leu Thr Val

225 230 235 240

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met

245 250 255

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser

260 265 270

Pro Gly Lys Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Cys

275 280 285

Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile

290 295 300

Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val

305 310 315 320

Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr

325 330 335

Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg

340 345 350

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

595 600 605

His His His His His

610

<210> 21

<211> 1839

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 21

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccgacaag acccacacct gccccccctg cccagcccca 180

gagctgctgg gcggaccctc cgtgttcctg ttcccccca agcccaagga caccctgatg 240

atcagcagga cccccgaggt gacctgcgtg gtggtggacg tgagccacga ggacccagag 300

gtgaagttca actggtacgt ggacggcgtg gaggtgcaca acgccaagac caagcccaga 360

gaggagcagt acaacagcac ctacagggtg gtgtccgtgc tgaccgtgct gcaccaggac 420

tggctgaacg gcaaggaata caagtgcaag gtctccaaca aggccctgcc agccccatc 480

gaaaagacca tcagcaaggc caagggccag ccacgggagc cccaggtgtg caccctgccc 540

ccctcccggg aggagatgac caagaaccag gtgtccctgt cctgtgcggt gaagggcttc 600

taccccagcg acatcgccgt ggagtgggag agcaacggcc agcccgagaa caactacaag 660

accacccccc cagtgctgga cagcgacggc agcttcttcc tggtcagcaa gctgaccgtg 720

gacaagtcca ggtggcagca gggcaacgtg ttcagctgca gcgtgatgca cgaggccctg 780

cacaaccact acacccagaa gagcctgagc ctgtcccccg gcaagggagg aagcggagga 840

tctggcggtt ccggaggctc ttgtgtggaa cccctcggca tggaaaacgg caatatcgcc 900

aatagccaga ttgccgccag cagcgtcaga gtgacatttc tgggactgca gcactgggtg 960

cccgagctgg ctagactgaa tagagccggc atggtcaacg cctggacacc cagcagcaac 1020

gacgataacc cttggattca agtgaacctg ctgcggcgta tgtgggtcac aggtgttgtt 1080

acacagggcg cctctagact ggccagccac gagtatctga aggcctttaa ggtggcctac 1140

agcctgaacg gccacgagtt cgacttcatc cacgacgtga acaagaagca caaagagttt 1200

gtcggcaact ggaacaagaa cgccgtgcac gtgaacctgt tcgagacacc tgtggaagcc 1260

cagtacgtgc ggctgtaccc tacaagctgt cacaccgcct gcactctgag attcgaactg 1320

ctgggatgcg agctgaacgg ctgtgctaat cctctgggcc tgaagaacaa cagcatcccc 1380

gataagcaga tcaccgccag ctccagctat aagacatggg gcctgcacct gttcagctgg 1440

aacccttctt acgccagact ggacaagcag ggcaacttca atgcttgggt ggccggcagc 1500

tacggcaatg atcagtggct gcaagtggac ctgggcagca gcaaagaagt gacaggcatc 1560

atcacccagg gcgccagaaa tttcggcagc gtgcagtttg tggccagcta caaagtggcc 1620

tactccaacg acagcgccaa ctggaccgag tatcaggacc ctagaaccgg cagctccaag 1680

atcttccccg gcaattggga caaccacagc cacaagaaga atctgttcga aacccctatc 1740

ctggccagat atgtgcgcat tctgcccgtg gcctggcaca acagaattgc cctgagactg 1800

gaactgctcg gctgtggctc tcaccaccac catcaccat 1839

<210> 22

<211> 596

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 22

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Cys Val
35 40 45

Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile Ala
50 55 60

Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val Pro
65 70 75 80

Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr Pro
85 90 95

Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg Arg
100 105 110
Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala Ser
115 120 125
His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly His
130 135 140
Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe Val
145 150 155 160
Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr Pro
165 170 175
Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr Ala
180 185 190
Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys Ala
195 200 205
Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile Thr
210 215 220
Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp Asn
225 230 235 240
Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp Val
245 250 255
Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly Ser
260 265 270
Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly
275 280 285
Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser
290 295 300
Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys Ile
305 310 315 320
Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe Glu
325 330 335
Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp His

340 345 350

Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys Gly Gly Gly Gly

355 360 365

Thr Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu

370 375 380

Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu

385 390 395 400

Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser

405 410 415

His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu

420 425 430

Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr

435 440 445

Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn

450 455 460

Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro

465 470 475 480

Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln

485 490 495

Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val

500 505 510

Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val

515 520 525

Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro

530 535 540

Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr

545 550 555 560

Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val

565 570 575

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu

580 585 590

Ser Pro Gly Lys

595

<210> 23

<211> 1788

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 23

ctggacatct gcagcaagaa cccctgccac aacggcggcc tgtgcgaaga gatcagccag 60

gaagtgcggg gcgacgtgtt ccccagctac acctgtacct gcctgaaggg ctacgccggc 120

aaccactgcg agactaagtg cgtggaaccc ctgggcatgg aaaacggcaa tatcgccaac 180

agccagatcg ccgccagctc cgtgcgcgtg acctttctgg gactgcagca ctgggtgccc 240

gagctggcca gactgaacag agccggcatg gtgaacgcct ggacccccag cagcaacgac 300

gacaaccctt ggatccaggt gaacctgctg cggcggatgt gggtgacagg cgtggtgaca 360

cagggcgcca gcagactggc cagccacgag tacctgaagg cctttaaggt ggcctacagc 420

ctgaacggcc acgagttcga cttcatccac gacgtgaaca agaaacacaa agaatttgtg 480

ggcaactgga acaagaacgc cgtgcacgtg aacctgttcg agacacccgt ggaagcccag 540

tacgtgcggc tgtaccccac cagctgccac accgcctgca ccctgagatt cgagctgctg 600

ggctgcgagc tgaacggctg cgccaacccc ctgggcctga agaacaacag catccccgac 660

aagcagatca ccgcctccag cagctacaag acctggggcc tgcacctgtt cagctggaac 720

cccagctacg cccggctgga caagcagggc aacttcaacg cctgggtggc cggcagctac 780

ggcaacgacc agtggctgca ggtggacctg ggcagcagca aagaagtgac cggcatcatc 840

acccaggggg ccagaaactt cggcagcgtg cagttcgtgg ccagctacaa agtggcctac 900

tccaacgaca gcgccaactg gaccgagtac caggaccccc ggaccggcag ctccaagatc 960

ttccccggca actgggacaa ccacagccac aagaagaatc tgttcgaaac ccccatcctg 1020

gccagatacg tgcggatcct gcccgtggcc tggcacaacc ggatcgccct gagactggaa 1080

ctgctgggat gtgggggagg cggtaccgac aagacccaca cctgcccccc ctgcccagcc 1140

ccagagctgc tgggcggacc ctccgtgttc ctgttccccc ccaagcccaa ggacaccctg 1200

atgatcagca ggacccccga ggtgacctgc gtggtggtgg acgtgagcca cgaggaccca 1260

gaggtgaagt tcaactggta cgtggacggc gtggaggtgc acaacgccaa gaccaagccc 1320

agagaggagc agtacaacag cacctacagg gtggtgtccg tgctgaccgt gctgcaccag 1380

gactggctga acggcaagga atacaagtgc aaggtctcca acaaggccct gccagccccc 1440

atcgaaaaga ccatcagcaa ggccaagggc cagccacggg agccccaggt gtacaccctg 1500

ccccccctcc gggaggagat gaccaagaac caggtgtccc tgacctgtct ggtgaagggc 1560

ttctacccca gcgacatcgc cgtggagtgg gagagcaacg gccagcccga gaacaactac 1620

aagaccaccc ccccagtgct ggacagcgac ggcagcttct tcctgtacag caagctgacc 1680

gtggacaagt ccaggtggca gcagggcaac gtgttcagct gcagcgtgat gcacgaggcc 1740

ctgcacaacc actacaccca gaagagcctg agcctgtccc ccggcaag 1788

<210> 24

<211> 604

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 24

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly

1 5 10 15

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met

20 25 30

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His

35 40 45

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val

50 55 60

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr

65 70 75 80

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly

85 90 95

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile

100 105 110

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

355 360 365

His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly His

370 375 380

Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe Val

385 390 395 400

Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr Pro

405 410 415

Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr Ala

420 425 430

Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys Ala

435 440 445

Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile Thr

450 455 460

Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp Asn

465 470 475 480

Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp Val

485 490 495

Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly Ser

500 505 510

Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly

515 520 525

Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser

530 535 540

Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys Ile

545 550 555 560

Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe Glu

565 570 575

Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp His

580 585 590

Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys

595 600

<210> 25

<211> 1812

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 25

gacaagaccc acacctgtcc cccctgccct gctcctgagc tgctgggagg acccagcgtg 60

ttcctgttcc cccccaagcc caaggacacc ctgatgatca gccggacccc cgaagtgacc 120

tgcgtggtgg tggacgtgtc ccacgaggac cctgaagtga agttcaattg gtacgtggac 180

ggcgtggagg tgcacaacgc caagaccaag ccccgggagg aacagtacaa cagcacctac 240

cgggtggtgt ccgtgctgac cgtgctgcac caggactggc tgaacggcaa agaatacaag 300

tgcaaggtgt ccaacaaggc cctgcctgcc cccatcgaga aaaccatcag caaggccaag 360

ggccagccca gagaacccca ggtgtacaca ctcccaccaa gccgggagga aatgaccaag 420

aaccaggtgt ccctgacctg cctggtgaag ggcttctacc ccagcgacat tgccgtggag 480

tgggagagca acggccagcc tgagaacaac tacaagacca cccctccagt cctcgattct 540

gatggatctt tcttcctgta ctccaagctg accgtggaca agagccggtg gcagcaggga 600

aacgtctttt cctgttccgt catgcatgag gctctccaca atcactacac ccagaagtcc 660

ctgagcctga gccccggcaa gggatccctc gaggtgctgt ttcagggacc aggcagcagc 720

ctggacatct gcagcaagaa cccctgccac aacggcggcc tgtgcgaaga gatcagccag 780

gaagtgcggg gcgacgtgtt ccccagctac acctgtacct gcctgaaggg ctacgccggc 840

aaccactgcg agactaagtg cgtggaaccc ctgggaatgg aaaacggcaa tatcgccaac 900

agccagatcg ccgccagctc cgtcagagtg acctttctgg gactccagca ctgggtgccc 960

gagctggcca gactgaatag agccggcatg gtcaacgcct ggacccccag cagcaacgac 1020

gacaacccct ggattcaagt gaacctgctg cggcgtatgt gggtcaccgg cgtcgtgaca 1080

cagggcgcta gcagactggc cagccacgag tacctgaagg cctttaaggt ggcctacagc 1140

ctgaacggcc acgagttcga cttcatccac gacgtgaaca agaaacacaa agaatttgtg 1200

ggcaactgga acaagaacgc cgtgcacgtg aacctgttcg agacacccgt ggaagcccag 1260

tacgtgcggc tgtaccctac cagctgtcac accgcctgca ccttaagatt cgagctgctg 1320

ggctgcgagc tgaacggctg cgctaatcct ctgggcctga agaacaacag catccccgac 1380

aagcagatca ccgcctccag cagctacaag acctggggac tgcacctgtt cagctggaac 1440

cctagctacg cccggctgga caagcagggc aacttcaatg cttgggtggc cggcagctac 1500

ggcaacgacc agtggctcca ggtggacctg ggcagcagca aagaagtgac cggcatcatc 1560

acccaggggg ccagaaactt cggcagcgtg cagttcgtgg cctcctacaa agtggcctac 1620

tccaacgaca gcgccaactg gaccgagtac caggacccta gaaccggcag ctccaagatt 1680

ttccccggca actgggataa ccacagccac aagaagaatc tgttcgaaac ccccatcctg 1740

gcccgctacg tgcgcattct accggtcgcc tggcacaacc ggatcgccct gagactggaa 1800

ctgctgggat gc 1812

<210> 26

<211> 364

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 26

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Cys
35 40 45

Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile
50 55 60

Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp
65 70 75 80

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

Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly
100 105 110

Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe

115 120 125

Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp

130 135 140

Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys

145 150 155 160

Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe

165 170 175

Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp

180 185 190

His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys Gly Cys Ala

195 200 205

Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile Thr

210 215 220

Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp Asn

225 230 235 240

Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp Val

245 250 255

Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly Ser

260 265 270

Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly

275 280 285

Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser

290 295 300

Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys Ile

305 310 315 320

Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe Glu

325 330 335

Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp His

340 345 350

Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys

355 360

<210> 27

<211> 1092

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 27

ctggacatct gcagcaagaa cccctgccac aacggcggcc tgtgcgaaga gatcagccag 60

gaagtgcggg gcgacgtgtt ccccagctac acctgtacct gcctgaaggg ctacgccggc 120

aaccactgcg agacaaaggg ctgcgccaac cccctgggcc tgaagaacaa cagcatcccc 180

gacaagcaga tcaccgccag cagcagctac aagacctggg gcctgcacct gttcagctgg 240

aaccccagct acgcccggct ggacaagcag ggcaacttca acgcctgggt ggccggcagc 300

tacggcaacg accagtggct gcaggtggac ctgggcagca gcaaagaagt gaccggcatc 360

atcacccagg gcgccagaaa cttcggcagc gtgcagttcg tggccagcta caaggtggcc 420

tacagcaacg acagcgccaa ctggaccgag taccaggacc cccggaccgg cagctccaag 480

atcttccccg gcaactggga caaccacagc cacaagaaga acctgttcga gacacccatc 540

ctggccagat acgtgcggat cctgcccgtg gcctggcaca accggatcgc cctgagactg 600

gaactgctgg gctgcggctg tgccaatcct ctgggactga aaaacaattc catccctgat 660

aagcagatta cagcctccag ctcctataag acatgggggc tgcatctgtt ttcttggaac 720

ccctcctacg ctagactgga taagcaggga aatttcaatg cttgggtggc cgggtcctat 780

ggaaatgatc agtggctgca ggtggacctg ggatcctcca aagaagtgac agggattatt 840

acacaggggg ctcggaactt tggctctgtg cagtttgtgg cttcctacaa agtggcttac 900

tccaacgatt ccgccaattg gacagaatat caggatccca gaaccggctc cagcaagatc 960

tttcctggaa attgggataa tcactcccac aagaaaaatc tgtttgaaac ccctattctg 1020

gctcgctatg tgcgcattct gcctgtggct tggcataata gaatcgctct gcggctggaa 1080

ctgctgggat gc 1092

<210> 28

<211> 967

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 28

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Cys Val
35 40 45

Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile Ala
50 55 60

Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val Pro
65 70 75 80

Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr Pro
85 90 95

Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg Arg
100 105 110

Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala Ser
115 120 125

His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly His
130 135 140

Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe Val
145 150 155 160

Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr Pro
165 170 175

Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr Ala
180 185 190

Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys Ala
195 200 205

Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile Thr

210 215 220

Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp Asn

225 230 235 240

Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp Val

245 250 255

Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly Ser

260 265 270

Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly

275 280 285

Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser

290 295 300

Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys Ile

305 310 315 320

Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe Glu

325 330 335

Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp His

340 345 350

Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys Gly Gly Ser Gly

355 360 365

Gly Ser Gly Gly Ser Gly Gly Ser Asp Ala His Lys Ser Glu Val Ala

370 375 380

His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu

385 390 395 400

Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val

405 410 415

Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp

420 425 430

Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp

435 440 445

Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala

450 455 460

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

705 710 715 720

Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys

725 730 735

Cys Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu

740 745 750

Phe Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys

755 760 765

Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu

770 775 780

Val Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val

785 790 795 800

Glu Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His

805 810 815

Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val

820 825 830

Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg

835 840 845

Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe

850 855 860

Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala

865 870 875 880

Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu

885 890 895

Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys

900 905 910

Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala

915 920 925

Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe

930 935 940

Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly

945 950 955 960

Leu His His His His His His

965

<210> 29

<211> 2901

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 29

ctggacatct gcagcaagaa cccctgccac aacggcggcc tgtgcgaaga gatcagccag 60

gaagtgcggg gcgacgtgtt ccccagctac acctgtacct gcctgaaggg ctacgccggc 120

aaccactgcg agactaagtg cgtggaaccc ctgggcatgg aaaacggcaa tatcgccaac 180

agccagatcg ccgccagctc cgtgcgcgtg acctttctgg gactgcagca ctgggtgccc 240

gagctggcca gactgaacag agccggcatg gtgaacgcct ggacccccag cagcaacgac 300

gacaaccctt ggatccaggt gaacctgctg cggcggatgt gggtgacagg cgtggtgaca 360

cagggcgcca gcagactggc cagccacgag tacctgaagg cctttaaggt ggcctacagc 420

ctgaacggcc acgagttcga cttcatccac gacgtgaaca agaaacacaa agaatttgtg 480

ggcaactgga acaagaacgc cgtgcacgtg aacctgttcg agacacccgt ggaagcccag 540

tacgtgcggc tgtaccccac cagctgccac accgcctgca ccctgagatt cgagctgctg 600

ggctgcgagc tgaacggctg cgccaacccc ctgggcctga agaacaacag catccccgac 660

aagcagatca ccgcctccag cagctacaag acctggggcc tgcacctgtt cagctggaac 720

cccagctacg cccggctgga caagcagggc aacttcaacg cctgggtggc cggcagctac 780

ggcaacgacc agtggctgca ggtggacctg ggcagcagca aagaagtgac cggcatcatc 840

acccaggggg ccagaaactt cggcagcgtg cagttcgtgg ccagctacaa agtggcctac 900

tccaacgaca gcgccaactg gaccgagtac caggaccccc ggaccggcag ctccaagatc 960

ttccccggca actgggacaa ccacagccac aagaagaatc tgttcgaaac ccccatcctg 1020

gccagatacg tgcggatcct gcccgtggcc tggcacaacc ggatcgccct gagactggaa 1080

ctgctgggat gtggaggaag cggaggatct ggcggttccg gaggctctga cgcccacaag 1140

agcgaggtgg cccaccggtt caaggacctg ggcgaggaaa acttcaaggc cctggtgctg 1200

atcgccttcg cccagtacct gcagcagagc cccttcgaag atcacgtaaa gttagtcaac 1260

gaggttacgg aattcgcaaa gacatgcgtt gctgacgaat ccgctgagaa ttgtgacaag 1320

agtttgcaca ctttattcgg agataagttg tgtactgtag ctactttgag agagacttac 1380

ggtgaaatgg ctgactgctg tgcaaaacag gaaccagaac gtaacgaatg tttccttcag 1440

cataaggatg ataaccctaa ccttccaagg cttgttaggc cagaagtcga cgtgatgtgc 1500

accgccttcc atgataatga agagactttt cttaaaaagt acctatacga gattgcaagg 1560

cgtcatccat atttttacgc cccagagctg ttgtttttcg caaagagata caaagctgca 1620

tttactgagt gttgccaagc tgccgacaag gccgcttgtt tgctaccaaa gttggacgaa 1680

ttgagagacg agggtaaggc atcatctgcc aagcagagat taaaatgtgc atctttgcaa 1740

aaatttggag agagagcttt taaggcatgg gctgttgccc gactaagcca aagattccca 1800

aaagccgaat ttgctgaagt atccaagctg gtgactgatt tgactaaagt acatacagaa 1860

tgttgccatg gcgacctttt agaatgtgct gatgacagag cagatttggc taagtatatc 1920

tgcgaaaatc aagattcaat cagctctaag ctgaaggaat gttgcgagaa accactgtta 1980

gaaaaatcgc attgtattgc tgaagttgaa aatgatgaga tgcctgctga cttgccttct 2040

cttgccgctg attttgttga gtcgaaggat gtctgtaaga attatgctga agctaaagac 2100

gttttcctgg gtatgttctt atatgagtac gcaagacgtc acccagatta ctctgtggtt 2160

ctgctactga gattggctaa aacatacgag acaacgctgg agaagtgctg tgctgccgct 2220

gaccctcatg agtgctatgc aaaggttttt gatgaattca aaccattggt tgaagagcct 2280

caaaacttga taaagcagaa ctgtgagctg tttgagcaat tgggtgagta taagttccaa 2340

aatgccctgt tggtgagata tacaaaaaag gtaccccaag tttcaacgcc cactttagtt 2400

gaagtgtcca gaaatcttgg taaagtgggt agcaaatgtt gcaagcatcc agaagccaag 2460

cgaatgccct gtgctgagga ttatctgtcc gtcgtgttga accaattgtg cgtattacac 2520

gaaaaaaccc cagtctctga tagagtcacc aaatgttgca ctgagtcact agttaataga 2580

aggccttgtt tttccgcttt ggaagttgat gaaacctacg tgcctaagga atttaacgct 2640

gagaccttta cctttcacgc tgacatttgt actttgagtg aaaaagagcg tcaaatcaaa 2700

aagcaaaccg ctcttgttga attggtgaaa cacaagccta aggctacgaa ggagcagctt 2760

aaagccgtca tggacgattt cgccgcattt gttgaaaaat gctgtaaagc tgatgacaag 2820

gaaacatgtt tcgctgaaga gggaaagaaa ttggttgcgg ccagtcaggc cgcacttggt 2880

ttgcaccatc atcaccatca c 2901

<210> 30

<211> 965

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 30

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1 5 10 15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
20 25 30

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
35 40 45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
50 55 60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65 70 75 80

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
85 90 95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
100 105 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
115 120 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
130 135 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145 150 155 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
165 170 175

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
180 185 190

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
195 200 205

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

450 455 460

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser

465 470 475 480

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr

485 490 495

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp

500 505 510

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala

515 520 525

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu

530 535 540

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys

545 550 555 560

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val

565 570 575

Ala Ala Ser Gln Ala Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly

580 585 590

Ser Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys

595 600 605

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

610 615 620

Cys Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Cys

625 630 635 640

Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile

645 650 655

Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val

660 665 670

Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr

675 680 685

Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg

690 695 700

Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala
705 710 715 720
Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly
725 730 735
His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe
740 745 750
Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr
755 760 765
Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr
770 775 780
Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys
785 790 795 800
Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile
805 810 815
Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp
820 825 830
Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp
835 840 845
Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly
850 855 860
Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe
865 870 875 880
Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp
885 890 895
Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys
900 905 910
Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe
915 920 925
Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp
930 935 940
His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys Gly Ser His

945 950 955 960

His His His His His

965

<210> 31

<211> 2895

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 31

gacgcccaca agagcgaggt ggcccaccgg ttcaaggacc tgggcgagga aaacttcaag 60

gccctggtgc tgatcgcctt cgcccagtac ctgcagcaga gccccttcga agatcacgta 120

aagttagtca acgaggttac ggaattcgca aagacatgcg ttgctgacga atccgctgag 180

aattgtgaca agagtttgca cactttattc ggagataagt tgtgtactgt agctactttg 240

agagagactt acggtgaaat ggctgactgc tgtgcaaaac aggaaccaga acgtaacgaa 300

tgtttccttc agcataagga tgataaccct aaccttccaa ggcttgttag gccagaagtc 360

gacgtgatgt gcaccgcctt ccatgataat gaagagactt ttcttaaaaa gtacctatac 420

gagattgcaa ggcgtcatcc atatttttac gccccagagc tgttgttttt cgcaaagaga 480

tacaaagctg catttactga gtgttgccaa gctgccgaca aggccgcttg tttgctacca 540

aagttggacg aattgagaga cgagggtaag gcatcatctg ccaagcagag attaaaatgt 600

gcatctttgc aaaaatttgg agagagagct tttaaggcat gggctgttgc ccgactaagc 660

caaagattcc caaaagccga atttgctgaa gtatccaagc tggtgactga tttgactaaa 720

gtacatacag aatgttgcca tggcgacctt ttagaatgtg ctgatgacag agcagatttg 780

gctaagtata tctgcgaaaa tcaagattca atcagctcta agctgaagga atgttgcgag 840

aaaccactgt tagaaaaatc gcattgtatt gctgaagttg aaaatgatga gatgcctgct 900

gacttgcctt ctcttgccgc tgattttgtt gagtcgaagg atgtctgtaa gaattatgct 960

gaagctaaag acgttttcct gggtatgttc ttatatgagt acgcaagacg tcacccagat 1020

tactctgtgg ttctgctact gagattggct aaaacatacg agacaacgct ggagaagtgc 1080

tgtgctgccg ctgaccctca tgagtgctat gcaaaggttt ttgatgaatt caaaccattg 1140

gttgaagagc ctcaaaactt gataaagcag aactgtgagc tgtttgagca attgggtgag 1200

tataagttcc aaaatgccct gttggtgaga tatacaaaaa aggtacccca agtttcaacg 1260

cccactttag ttgaagtgtc cagaaatctt ggtaaagtgg gtagcaaatg ttgcaagcat 1320

ccagaagcca agcgaatgcc ctgtgctgag gattatctgt ccgtcgtgtt gaaccaattg 1380

tgcgtattac acgaaaaaac cccagtctct gatagagtca ccaaatgttg cactgagtca 1440

ctagttaata gaaggccttg tttttccgct ttggaagttg atgaaaccta cgtgcctaag 1500

gaatttaacg ctgagacctt tacctttcac gctgacattt gtactttgag tgaaaaagag 1560

cgtcaaatca aaaagcaaac cgctcttgtt gaattggtga aacacaagcc taaggctacg 1620

aaggagcagc ttaaagccgt catggacgat ttcgccgcat ttgttgaaaa atgctgtaaa 1680

gctgatgaca aggaaacatg tttcgctgaa gagggaaaga aattggttgc ggccagtcag 1740

gccggaggaa gcggaggatc tggcggttcc ggaggctctc tagacatctg cagcaagaac 1800

ccctgccaca acggcggcct gtgcgaagag atcagccagg aagtgcgggg cgacgtgttc 1860

cccagctaca cctgtacctg cctgaagggc tacgccggca accactgcga gactaagtgc 1920

gtggaacccc tgggcatgga aaacggcaat atcgccaaca gccagatcgc cgccagctcc 1980

gtgcgcgtga cctttctggg actgcagcac tgggtgcccg agctggccag actgaacaga 2040

gccggcatgg tgaacgcctg gacccccagc agcaacgacg acaacccttg gatccaggtg 2100

aacctgctgc ggcggatgtg ggtgacaggc gtggtgacac agggcgccag cagactggcc 2160

agccacgagt acctgaaggc ctttaaggtg gcctacagcc tgaacggcca cgagttcgac 2220

ttcatccacg acgtgaacaa gaaacacaaa gaatttgtgg gcaactggaa caagaacgcc 2280

gtgcacgtga acctgttcga gacacccgtg gaagcccagt acgtgcggct gtaccccacc 2340

agctgccaca ccgcctgcac cctgagattc gagctgctgg gctgcgagct gaacggctgc 2400

gccaaccccc tgggcctgaa gaacaacagc atccccgaca agcagatcac cgcctccagc 2460

agctacaaga cctggggcct gcacctgttc agctggaacc ccagctacgc ccggctggac 2520

aagcagggca acttcaacgc ctgggtggcc ggcagctacg gcaacgacca gtggctgcag 2580

gtggacctgg gcagcagcaa agaagtgacc ggcatcatca cccaggggc cagaaacttc 2640

ggcagcgtgc agttcgtggc cagctacaaa gtggcctact ccaacgacag cgccaactgg 2700

accgagtacc aggacccccg gaccggcagc tccaagatct tccccggcaa ctgggacaac 2760

cacagccaca agaagaatct gttcgaaacc cccatcctgg ccagatacgt gcggatcctg 2820

cccgtggcct ggcacaaccg gatcgccctg agactggaac tgctgggatg tggctctcac 2880

caccaccatc accat 2895

<210> 32

<211> 641

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 32

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
100 105 110

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
115 120 125

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
130 135 140

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
145 150 155 160

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
165 170 175

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
180 185 190

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg

195 200 205

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala

210 215 220

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser

225 230 235 240

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu

245 250 255

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro

260 265 270

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

275 280 285

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp

290 295 300

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser

305 310 315 320

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His

325 330 335

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

340 345 350

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

355 360 365

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg

370 375 380

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr

385 390 395 400

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu

405 410 415

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

420 425 430

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu

435 440 445

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

<210> 33
<211> 1923
<212> ADN
<213> Trình tự Nhân tạo
<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 33

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccgatgct cacaagagcg aggtggccca cagattcaag 180

gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240

cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300

tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360

aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420

aagcaagagc ccgagagaaa cgagtgcttc ctgcagcaca aggacgacaa ccccaacctg 480

cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540

accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgcccct 600

gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa gtgcgccagc ctgcagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgtaccc tgagcgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaggctgct ctcggacttg gctctcacca ccaccatcac 1920

cat 1923

<210> 34

<211> 809

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 34

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
100 105 110

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
115 120 125

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
130 135 140

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

385 390 395 400

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu

405 410 415

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

420 425 430

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu

435 440 445

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro

450 455 460

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys

465 470 475 480

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys

485 490 495

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His

500 505 510

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser

515 520 525

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr

530 535 540

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp

545 550 555 560

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala

565 570 575

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu

580 585 590

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys

595 600 605

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val

610 615 620

Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly

625 630 635 640

Gly Ser Gly Gly Ser Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn
645 650 655

Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu
660 665 670

Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly
675 680 685

Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile
690 695 700

Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr Gly Val Val Thr Gln
705 710 715 720

Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val
725 730 735

Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn
740 745 750

Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His
755 760 765

Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr
770 775 780

Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly
785 790 795 800

Cys Gly Ser His His His His His His
805

<210> 35

<211> 2427

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 35

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccgatgct cacaagagcg aggtggccca cagattcaag 180

gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240

cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300

tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360

aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420

aagcaagagc ccgagagaaa cgagtgcttc ctgcagcaca aggacgacaa ccccaacctg 480

cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540

accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgcccct 600

gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa gtgcgccagc ctgcagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgtaccc tgagcgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaggctgct ctcggacttg gaggaagcgg aggatctggc 1920

ggttccggag gctcttgtgt ggaacccctc ggcatggaaa acggcaatat cgccaatagc 1980

cagattgccg ccagcagcgt cagagtgaca tttctgggac tgcagcactg ggtgcccgag 2040

ctggctagac tgaatagagc cggcatggtc aacgcctgga cacccagcag caacgacgat 2100

aacccttgga ttcaagtgaa cctgctgcgg cgtatgtggg tcacaggtgt tgttacacag 2160

ggcgcctcta gactggccag ccacgagtat ctgaaggcct ttaaggtggc ctacagcctg 2220

aacggccacg agttcgactt catccacgac gtgaacaaga agcacaaaga gtttgtcggc 2280

aactggaaca agaacgccgt gcacgtgaac ctgttcgaga cacctgtgga agcccagtac 2340

gtgcggctgt accctacaag ctgtcacacc gcctgcactc tgagattcga actgctggga 2400

tgcggctctc accaccacca tcaccat 2427

<210> 36

<211> 811

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 36

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu

1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys

20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser

35 40 45

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

50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln

65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu

85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys

100 105 110

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

355 360 365

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg

370 375 380

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr

385 390 395 400

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu

405 410 415

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

420 425 430

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu

435 440 445

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro

450 455 460

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys

465 470 475 480

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys

485 490 495

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His

500 505 510

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser

515 520 525

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr

530 535 540

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp

545 550 555 560

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala

565 570 575

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu

580 585 590

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys

595 600 605

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
610 615 620
Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly
625 630 635 640
Gly Ser Gly Gly Ser Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser
645 650 655
Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly
660 665 670
Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln
675 680 685
Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp
690 695 700
Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr
705 710 715 720
Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys
725 730 735
Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro
740 745 750
Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser
755 760 765
His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg
770 775 780
Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu
785 790 795 800
Leu Gly Cys Gly Ser His His His His His His
805 810

<210> 37

<211> 2433

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 37

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccgacgcc cacaagagcg aggtggccca ccggttcaag 180

gacctgggcg aggaaaactt caaggccctg gtgctgatcg ccttcgccca gtacctgcag 240

cagagcccct tcgaagatca cgtaaagtta gtcaacgagg ttacggaatt cgcaaagaca 300

tgcgttgctg acgaatccgc tgagaattgt gacaagagtt tgcacacttt attcggagat 360

aagttgtgta ctgtagctac tttgagagag acttacggtg aaatggctga ctgctgtgca 420

aaacaggaac cagaacgtaa cgaatgtttc cttcagcata aggatgataa ccctaacctt 480

ccaaggcttg ttaggccaga agtcgacgtg atgtgcaccg ccttccatga taatgaagag 540

acttttctta aaaagtacct atacgagatt gcaaggcgtc atccatattt ttacgcccca 600

gagctgttgt ttttcgcaaa gagatacaaa gctgcattta ctgagtgttg ccaagctgcc 660

gacaaggccg cttgtttgct accaaagttg gacgaattga gagacgaggg taaggcatca 720

tctgccaagc agagattaaa atgtgcatct ttgcaaaaat ttggagagag agcttttaag 780

gcatgggctg ttgcccgact aagccaaaga ttcccaaaag ccgaatttgc tgaagtatcc 840

aagctggtga ctgatttgac taaagtacat acagaatgtt gccatggcga ccttttagaa 900

tgtgctgatg acagagcaga tttggctaag tatatctgcg aaaatcaaga ttcaatcagc 960

tctaagctga aggaatgttg cgagaaacca ctgttagaaa aatcgcattg tattgctgaa 1020

gttgaaaatg atgagatgcc tgctgacttg ccttctcttg ccgctgattt tgttgagtcg 1080

aaggatgtct gtaagaatta tgctgaagct aaagacgttt tcctgggtat gttcttatat 1140

gagtacgcaa gacgtcaccc agattactct gtggttctgc tactgagatt ggctaaaaca 1200

tacgagacaa cgctggagaa gtgctgtgct gccgctgacc ctcatgagtg ctatgcaaag 1260

gtttttgatg aattcaaacc attggttgaa gagcctcaaa acttgataaa gcagaactgt 1320

gagctgtttg agcaattggg tgagtataag ttccaaaatg ccctgttggt gagatataca 1380

aaaaaggtac cccaagtttc aacgcccact ttagttgaag tgtccagaaa tcttggtaaa 1440

gtgggtagca aatgttgcaa gcatccagaa gccaagcgaa tgccctgtgc tgaggattat 1500

ctgtccgtcg tgttgaacca attgtgcgta ttacacgaaa aaaccccagt ctctgataga 1560

gtcaccaaat gttgcactga gtcactagtt aatagaaggc cttgtttttc cgctttggaa 1620

gttgatgaaa cctacgtgcc taaggaattt aacgctgaga cctttacctt tcacgctgac 1680

atttgtactt tgagtgaaaa agagcgtcaa atcaaaaagc aaaccgctct tgttgaattg 1740

gtgaaacaca agcctaaggc tacgaaggag cagcttaaag ccgtcatgga cgatttcgcc 1800

gcatttgttg aaaaatgctg taaagctgat gacaaggaaa catgtttcgc tgaagaggga 1860

aagaaattgg ttgcggccag tcaggccgca cttggtttgg gaggaagcgg aggatctggc 1920

ggttccggag gctcttgcgc caacccccctg ggcctgaaga acaacagcat ccccgacaag 1980

cagatcaccg cctccagcag ctacaagacc tggggcctgc acctgttcag ctggaacccc 2040

agctacgccc ggctggacaa gcagggcaac ttcaacgcct gggtggccgg cagctacggc 2100

aacgaccagt ggctgcaggt ggacctgggc agcagcaaag aagtgaccgg catcatcacc 2160

caggggggca gaaacttcgg cagcgtgcag ttcgtggcca gctacaaagt ggcctactcc 2220

aacgacagcg ccaactggac cgagtaccag gacccccgga ccggcagctc caagatcttc 2280

cccggcaact gggacaacca cagccacaag aagaatctgt tcgaaacccc catcctggcc 2340

agatacgtgc ggatcctgcc cgtggcctgg cacaaccgga tcgccctgag actggaactg 2400

ctgggatgtg gctctcacca ccaccatcac cat 2433

<210> 38

<211> 971

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 38

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Glu Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

```
Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
                85                  90                  95
Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
            100                 105                 110
Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
        115                 120                 125
Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
    130                 135                 140
Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
145                 150                 155                 160
Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
                165                 170                 175
Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
            180                 185                 190
Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
        195                 200                 205
Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
    210                 215                 220
Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
225                 230                 235                 240
Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
                245                 250                 255
Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
            260                 265                 270
Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
        275                 280                 285
Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
    290                 295                 300
Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
305                 310                 315                 320
Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
```

325 330 335

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

340 345 350

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

355 360 365

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg

370 375 380

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr

385 390 395 400

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu

405 410 415

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

420 425 430

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu

435 440 445

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro

450 455 460

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys

465 470 475 480

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys

485 490 495

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His

500 505 510

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser

515 520 525

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr

530 535 540

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp

545 550 555 560

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala

565 570 575

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
580 585 590
Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
595 600 605
Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
610 615 620
Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly
625 630 635 640
Gly Ser Gly Gly Ser Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn
645 650 655
Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu
660 665 670
Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly
675 680 685
Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile
690 695 700
Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr Gly Val Val Thr Gln
705 710 715 720
Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val
725 730 735
Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn
740 745 750
Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His
755 760 765
Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr
770 775 780
Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly
785 790 795 800
Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser
805 810 815
Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly

820 825 830

Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln

835 840 845

Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp

850 855 860

Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr

865 870 875 880

Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys

885 890 895

Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro

900 905 910

Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser

915 920 925

His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg

930 935 940

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

945 950 955 960

Leu Gly Cys Gly Ser His His His His His His

965 970

<210> 39

<211> 2913

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 39

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgaggtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccgatgct cacaagagcg aggtggccca cagattcaag 180

gatctgggcg acgagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240

cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300

tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360

aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420

aagcaagagc ccgagagaaa cgagtgcttc ctgcagcaca aggacgacaa ccccaacctg 480

cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540

accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgcccct 600

gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa gtgcgccagc ctgcagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgtaccc tgagcgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaggctgct ctcggacttg gaggaagcgg aggatctggc 1920

ggttccggag gctcttgtgt ggaacccctc ggcatggaaa acggcaatat cgccaatagc 1980

cagattgccg ccagcagcgt cagagtgaca tttctgggac tgcagcactg ggtgcccgag 2040

ctggctagac tgaatagagc cggcatggtc aacgcctgga cacccagcag caacgacgat 2100

aacccttgga ttcaagtgaa cctgctgcgg cgtatgtggg tcacaggtgt tgttacacag 2160

ggcgcctcta gactggccag ccacgagtat ctgaaggcct ttaaggtggc ctacagcctg 2220

aacggccacg agttcgactt catccacgac gtgaacaaga agcacaaaga gtttgtcggc 2280

aactggaaca agaacgccgt gcacgtgaac ctgttcgaga cacctgtgga agcccagtac 2340

gtgcggctgt accctacaag ctgtcacacc gcctgcactc tgagattcga actgctggga 2400

tgcgagctga acggctgtgc taatcctctg ggcctgaaga acaacagcat ccccgataag 2460

cagatcaccg ccagctccag ctataagaca tggggcctgc acctgttcag ctggaaccct 2520

tcttacgcca gactggacaa gcagggcaac ttcaatgctt gggtggccgg cagctacggc 2580

aatgatcagt ggctgcaagt ggacctgggc agcagcaaag aagtgacagg catcatcacc 2640

cagggcgcca gaaatttcgg cagcgtgcag tttgtggcca gctacaaagt ggcctactcc 2700

aacgacagcg ccaactggac cgagtatcag gaccctagaa ccggcagctc caagatcttc 2760

cccggcaatt gggacaacca cagccacaag aagaatctgt tcgaaacccc tatcctggcc 2820

agatatgtgc gcattctgcc cgtggcctgg cacaacagaa ttgccctgag actggaactg 2880

ctcggctgtg gctctcacca ccaccatcac cat 2913

<210> 40

<211> 592

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 40

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe
50 55 60

Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr

65 70 75 80

Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser

85 90 95

Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala

100 105 110

Lys Arg Met Pro Cys Ala Glu Asp Cys Leu Ser Val Phe Leu Asn Gln

115 120 125

Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys

130 135 140

Cys Cys Thr Glu Ser Leu Val Asn Gly Arg Pro Cys Phe Ser Ala Leu

145 150 155 160

Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe

165 170 175

Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile

180 185 190

Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala

195 200 205

Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val

210 215 220

Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu

225 230 235 240

Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly

245 250 255

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Cys Val Glu Pro Leu Gly

260 265 270

Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val

275 280 285

Arg Val Thr Phe Leu Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg

290 295 300

Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp

305 310 315 320

Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr
325 330 335
Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu
340 345 350
Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe
355 360 365
Ile His Asp Val Asn Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn
370 375 380
Lys Asn Ala Val His Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln
385 390 395 400
Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg
405 410 415
Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly
420 425 430
Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser
435 440 445
Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala
450 455 460
Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr
465 470 475 480
Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val
485 490 495
Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe
500 505 510
Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr
515 520 525
Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn
530 535 540
Trp Asp Asn His Ser His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu
545 550 555 560
Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala

565 570 575

Leu Arg Leu Glu Leu Leu Gly Cys Gly Ser His His His His His His

580 585 590

<210> 41

<211> 1776

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 41

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccctggtg gaagaacccc agaacctgat caagcagaac 180

tgcgagctgt tcgagcagct gggcgagtac aagttccaga atgccctgct cgtgcggtac 240

accaagaaag tgcctcaggt gtccacacct acactggtcg aggtgtccag aaacctgggc 300

aaagtgggca gcaagtgctg caagcaccct gaggccaaaa gaatgccttg cgccgaggat 360

tgcctgagcg tgttcctgaa tcagctgtgc gtgctgcacg agaaaacccc tgtgtccgac 420

agagtgacca agtgctgtac cgagagcctg gtcaacggca gaccttgctt tagcgccctg 480

gaagtggatg agacatacgt gcccaaagag ttcaacgccg agacattcac cttccacgcc 540

gacatctgta ccctgagcga gaaagagcgg cagatcaaga agcagacagc cctggtcgag 600

ctggtcaagc acaagcctaa ggccaccaaa gaacagctga aggccgtgat ggacgacttc 660

gccgccttcg tggaaaagtg ttgcaaggcc gacgacaaag agacatgctt cgccgaagag 720

ggcaagaaac tggtggctgc ctctcaggct gctctcggac ttggaggaag cggaggatct 780

ggcggttccg gaggctcttg tgtggaaccc ctcggcatgg aaaacggcaa tatcgccaat 840

agccagatcg ccgccagcag cgtcagagtg acatttctgg gactgcagca ctgggtgcca 900

gagctggcta gactgaatag agccggcatg gtcaacgcct ggacacccag cagcaacgac 960

gacaacccct ggattcaagt gaacctgctg cggcgtatgt gggtcacagg tgttgttaca 1020

cagggcgcct ctagactggc cagccacgag tatctgaagg cctttaaggt ggcctacagc 1080

ctgaacggcc acgagttcga cttcatccac gacgtgaaca agaagcacaa agagtttgtc 1140

ggcaactgga acaagaacgc cgtgcacgtg aacctgttcg agacacctgt ggaagcccag 1200

tacgtgcggc tgtaccctac aagctgtcac accgcctgca cactgagatt cgaactgctg 1260

ggatgcgagc tgaacggctg tgctaatcct ctgggcctga agaacaacag catccccgat 1320

aagcagatca ccgcctccag cagctataag acatggggcc tgcacctgtt cagctggaac 1380

cctagctacg ccagactgga caagcagggc aactttaatg cctgggtggc cggcagctac 1440

ggcaatgatc aatggctgca agtggacctg ggcagcagca aagaagtgac cggcatcatt 1500

acccagggcg caagaaattt cggcagcgtg cagttcgtgg ccagctacaa agtggcctac 1560

tccaacgaca gcgccaactg gaccgagtat caggacccta gaaccggcag ctccaagatc 1620

ttccccggca attgggacaa ccacagccac aagaagaatc tgttcgaaac ccctatcctg 1680

gccagatatg tgcgcattct gcccgtggcc tggcacaaca gaattgccct gagactggaa 1740

ctgctcggct gtggctctca ccaccaccat caccat 1776

<210> 42

<211> 963

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 42

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys

100 105 110

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu

115 120 125

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro

130 135 140

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu

145 150 155 160

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His

165 170 175

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg

180 185 190

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg

195 200 205

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala

210 215 220

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser

225 230 235 240

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu

245 250 255

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro

260 265 270

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

275 280 285

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp

290 295 300

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser

305 310 315 320

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His

325 330 335

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

340 345 350

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

595 600 605

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val

610 615 620

Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly

625 630 635 640

Gly Ser Gly Gly Ser Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn

645 650 655

Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu

660 665 670

Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly

675 680 685

Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile

690 695 700

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

705 710 715 720

Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val

725 730 735

Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn

740 745 750

Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His

755 760 765

Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr

770 775 780

Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly

785 790 795 800

Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser

805 810 815

Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly

820 825 830

Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln

835 840 845

Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp
850 855 860
Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr
865 870 875 880
Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys
885 890 895
Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro
900 905 910
Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser
915 920 925
His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg
930 935 940
Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu
945 950 955 960
Leu Gly Cys

<210> 43
<211> 2889
<212> ADN
<213> Trình tự nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"
<400> 43
ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60
gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120
aaccactgcg agacaaaagg atccgatgct cacaagagcg aggtggccca cagattcaag 180
gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240
cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300
tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360
aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420

aagcaagagc ccgagagaaa cgagtgcttc ctgcagcaca aggacgacaa ccccaacctg 480

cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540

accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgccccc 600

gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa gtgcgccagc ctgcagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgtaccc tgagcgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaggctgct ctcggacttg gaggaagcgg aggatctggc 1920

ggttccggag gctcttgtgt ggaacccctc ggcatggaaa acggcaatat cgccaatagc 1980

cagattgccg ccagcagcgt cagagtgaca tttctgggac tgcagcactg ggtgcccgag 2040

ctggctagac tgaatagagc cggcatggtc aacgcctgga cacccagcag caacgacgat 2100

aacccttgga ttcaagtgaa cctgctgcgg cgtatgtggg tcacaggtgt tgttacacag 2160

ggcgcctcta gactggccag ccacgagtat ctgaaggcct ttaaggtggc ctacagcctg 2220

aacggccacg agttcgactt catccacgac gtgaacaaga agcacaaaga gtttgtcggc 2280

aactggaaca agaacgccgt gcacgtgaac ctgttcgaga cacctgtgga agcccagtac 2340

gtgcggctgt accctacaag ctgtcacacc gcctgcactc tgagattcga actgctggga 2400

tgcgagctga acggctgtgc taatcctctg ggcctgaaga acaacagcat ccccgataag 2460

cagatcaccg ccagctccag ctataagaca tggggcctgc acctgttcag ctggaaccct 2520

tcttacgcca gactggacaa gcagggcaac ttcaatgctt gggtggccgg cagctacggc 2580

aatgatcagt ggctgcaagt ggacctgggc agcagcaaag aagtgacagg catcatcacc 2640

cagggcgcca gaaatttcgg cagcgtgcag tttgtggcca gctacaaagt ggcctactcc 2700

aacgacagcg ccaactggac cgagtatcag gaccctagaa ccggcagctc caagatcttc 2760

cccggcaatt gggacaacca cagccacaag aagaatctgt tcgaaacccc tatcctggcc 2820

agatatgtgc gcattctgcc cgtggcctgg cacaacagaa ttgccctgag actggaactg 2880

ctcggctgt 2889

<210> 44

<211> 971

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 44

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
100 105 110
Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
115 120 125
Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
130 135 140
Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
145 150 155 160
Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
165 170 175
Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
180 185 190
Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
195 200 205
Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
210 215 220
Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
225 230 235 240
Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
245 250 255
Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
260 265 270
Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
275 280 285
Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
290 295 300
Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
305 310 315 320
Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
325 330 335
Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

340 345 350

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

355 360 365

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg

370 375 380

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr

385 390 395 400

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu

405 410 415

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

420 425 430

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu

435 440 445

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro

450 455 460

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys

465 470 475 480

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys

485 490 495

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His

500 505 510

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser

515 520 525

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr

530 535 540

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp

545 550 555 560

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala

565 570 575

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu

580 585 590

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
595 600 605

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
610 615 620

Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly
625 630 635 640

Gly Ser Gly Gly Ser Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn
645 650 655

Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu
660 665 670

Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly
675 680 685

Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile
690 695 700

Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr Gly Val Val Thr Gln
705 710 715 720

Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val
725 730 735

Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn
740 745 750

Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His
755 760 765

Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr
770 775 780

Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly
785 790 795 800

Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser
835 810 815

Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly
820 825 830

Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln

835 840 845

Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp

850 855 860

Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr

865 870 875 880

Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys

885 890 895

Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro

900 905 910

Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser

915 920 925

His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg

930 935 940

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

945 950 955 960

Leu Gly Cys Gly Ser His His His His His His

965 970

<210> 45

<211> 2913

<212> ADN

<213> Trình tự nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 45

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaagg atccgatgct cacaagagcg aggtggccca cagattcaag 180

gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240

cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300

tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360

aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420

aagcaagagc ccgagagaaa cgagtgcttc ctgcagcaca aggacgacaa ccccaacctg 480

cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540

accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgcccct 600

gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa gtgcgccagc ctgcagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgtaccc tgagcgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaggctgct ctcggacttg gaggaagcgg aggatctggc 1920

ggttccggag gctcttgtgt ggaacccctc ggcatggaaa acggcaatat cgccaatagc 1980

cagattgccg ccagcagcgt cagagtgaca tttctgggac tgcagcactg ggtgcccgag 2040

ctggctagac tgaatagagc cggcatggtc aacgcctgga cacccagcag caacgacgat 2100

aacccttgga ttcaagtgaa cctgctgcgg cgtatgtggg tcacaggtgt tgttacacag 2160

ggcgcctcta gactggccag ccacgagtat ctgaaggcct ttaaggtggc ctacagcctg 2220

aacggccacg agttcgactt catccacgac gtgaacaaga agcacaaaga gtttgtcggc 2280

aactggaaca agaacgccgt gcacgtgaac ctgttcgaga cacctgtgga agcccagtac 2340

gtgcggctgt accctacaag ctgtcacacc gcctgcactc tgagattcga actgctggga 2400

tgcgagctga acggctgtgc taatcctctg ggcctgaaga acaacagcat ccccgataag 2460

cagatcaccg ccagctccag ctataagaca tggggcctgc acctgttcag ctggaaccct 2520

tcttacgcca gactggacaa gcagggcaac ttcaatgctt gggtggccgg cagctacggc 2580

aatgatcagt ggctgcaagt ggacctgggc agcagcaaag aagtgacagg catcatcacc 2640

cagggcgcca gaaatttcgg cagcgtgcag tttgtggcca gctacaaagt ggcctactcc 2700

aacgacagcg ccaactggac cgagtatcag gaccctagaa ccggcagctc caagatcttc 2760

cccggcaatt gggacaacca cagccacaag aagaatctgt tcgaaacccc tatcctggcc 2820

agatatgtgc gcattctgcc cgtggcctgg cacaacagaa ttgccctgag actggaactg 2880

ctcggctgtg gctctcacca ccaccatcac cat 2913

<210> 46

<211> 947

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 46

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu

1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys

20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala

35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn

50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser

65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala

85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu

100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu

115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg

130 135 140

Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg

145 150 155 160

Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn

165 170 175

Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His

180 185 190

Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys

195 200 205

Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu

210 215 220

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

225 230 235 240

Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala

245 250 255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala

260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His

275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala

290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys

305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile

325 330 335

Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala
340 345 350
Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala
355 360 365
Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His
370 375 380
Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu
385 390 395 400
Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr
405 410 415
Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn
420 425 430
Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys
435 440 445
Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val
450 455 460
Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly
465 470 475 480
Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu
485 490 495
Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys
500 505 510
Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val
515 520 525
Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val
530 535 540
Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys
545 550 555 560
Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val
565 570 575
Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala

580 585 590

Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp

595 600 605

Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala

610 615 620

Ser Gln Ala Ala Leu Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn

625 630 635 640

Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu

645 650 655

Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly

660 665 670

Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile

675 680 685

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

690 695 700

Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val

705 710 715 720

Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn

725 730 735

Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His

740 745 750

Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr

755 760 765

Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly

770 775 780

Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser

785 790 795 800

Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly

805 810 815

Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln

820 825 830

Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp
835 840 845

Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr
850 855 860

Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys
865 870 875 880

Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro
885 890 895

Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser
900 905 910

His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg
915 920 925

Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu
930 935 940

Leu Gly Cys
945

<210> 47

<211> 2841

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 47

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaagga tgcccacaag agcgaggtgg cccacagatt caaggatctg 180

ggcgaagaga acttcaaggc cctggtgctg atcgccttcg ctcagtatct ccagcagagc 240

cctttcgagg accacgtgaa gctggtcaac gaagtgaccg agttcgccaa gacctgtgtg 300

gccgatgaga gcgccgagaa ctgtgataag agcctgcaca ccctgttcgg cgacaagctg 360

tgtacagtgg ccacactgag agaaacctac ggcgagatgg ccgactgctg tgccaagcaa 420

gagcccgaga gaaacgagtg cttcctccag cacaaggacg acaaccccaa cctgcctaga 480

ctcgtgcgac ccgaagtgga tgtgatgtgc accgcctttc acgacaacga ggaaaccttc 540

ctgaagaagt acctgtacga gatcgccaga cggcacccct acttttatgc ccctgagctg 600

ctgttcttcg ccaagcggta taaggccgcc ttcaccgaat gttgccaggc cgctgataag 660

gctgcctgtc tgctgcctaa gctggacgag ctgagagatg agggcaaagc cagctctgcc 720

aagcagagac tgaaatgcgc cagcctccag aagttcggcg agagagcttt taaggcctgg 780

gccgttgcca gactgagcca gagatttcct aaggccgagt ttgccgaggt gtccaagctc 840

gtgaccgatc tgacaaaggt gcacaccgag tgctgtcacg gcgatctgct ggaatgtgcc 900

gacgatagag ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 960

ctgaaagagt gctgcgagaa gcccctgctg gaaaagtctc actgtatcgc cgaggtggaa 1020

aacgacgaga tgcctgccga tctgcctagc ctggctgccg atttcgtgga aagcaaggac 1080

gtgtgcaaga actacgccga ggccaaggat gtgtttctgg gcatgtttct gtatgagtac 1140

gcccgcagac accccgacta ttctgtggtt ctgctgctgc ggctggccaa aacctacgag 1200

acaaccctgg aaaaatgctg cgccgctgcc gatcctcacg agtgttatgc caaggtgttc 1260

gacgagttca agcctctggt ggaagaaccc cagaacctga tcaagcagaa ctgcgagctg 1320

ttcgagcagc tgggcgagta caagttccag aatgccctgc tcgtgcggta caccaagaaa 1380

gtgcctcagg tgtccacacc tacactggtt gaggtgtccc ggaatctggg caaagtgggc 1440

agcaagtgtt gcaagcaccc tgaggccaag agaatgcctt gcgccgagga ttacctgagc 1500

gtggtgctga atcagctgtg cgtgctgcac gagaaaaccc ctgtgtccga cagagtgacc 1560

aagtgctgta ccgagagcct cgtgaacaga aggccttgct ttagcgccct ggaagtggac 1620

gagacatacg tgcccaaaga gttcaacgcc gagacattca ccttccacgc cgacatctgc 1680

accctgtccg agaaagagcg gcagatcaag aagcagacag ccctggtcga gctggttaag 1740

cacaagccca aggccaccaa agaacagctg aaggccgtga tggacgactt cgccgccttt 1800

gtcgagaagt gctgcaaggc cgacgacaaa gagacatgct tcgccgaaga gggcaagaaa 1860

ctggtggctg cttctcaggc cgctctgtgt gtggaacccc tcggcatgga aaacggcaat 1920

atcgccaata gccagattgc cgccagcagc gtcagagtga catttctggg actgcaacac 1980

tgggtgcccg agctggctag actgaataga gccggcatgg tcaacgcctg gacacccagc 2040

agcaacgacg ataatccctg gattcaagtg aacctgctgc ggcgtatgtg ggtcacaggt 2100

gttgttacac agggcgcaag cagactggcc agccacgagt atctgaaggc ctttaaggtg 2160

gcctacagcc tgaacggcca cgagttcgac ttcatccacg acgtgaacaa gaagcacaaa 2220

gagtttgtcg gcaactggaa caagaacgcc gtgcacgtga acctgttcga gacacctgtg 2280

gaagcccagt acgtgcggct gtaccctaca agctgtcaca ccgcctgcac tctgagattc 2340

gaactgctgg gatgcgagct gaacggctgt gctaatcctc tgggcctgaa gaacaacagc 2400

atccccgata agcagatcac cgccagctcc agctataaga catggggcct gcacctgttc 2460

agctggaacc cttcttacgc cagactggac aagcagggca acttcaatgc ttgggtggcc 2520

ggcagctacg gcaatgatca gtggctgcaa gtggacctgg gcagcagcaa agaagtgaca 2580

ggcatcatca cccaaggggc cagaaatttc ggcagcgtgc agttcgtggc cagctacaaa 2640

gtggcctact ccaacgacag cgccaactgg accgagtatc aggaccctag aaccggcagc 2700

tccaagatct tccccggcaa ttgggacaac cacagccaca agaagaatct gttcgaaacc 2760

cctatcctgg ccagatatgt gcgcattctg cccgtggcct ggcacaacag aattgccctg 2820

agactggaac tgctcggctg c 2841

<210> 48

<211> 941

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 48

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala
35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn
50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser
65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala
85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu

100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu

115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg

130 135 140

Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg

145 150 155 160

Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn

165 170 175

Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His

180 185 190

Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys

195 200 205

Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu

210 215 220

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

225 230 235 240

Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala

245 250 255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala

260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His

275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala

290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys

305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile

325 330 335

Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala

340 345 350

Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala
355 360 365

Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His
370 375 380

Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu
385 390 395 400

Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr
405 410 415

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

Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys
435 440 445

Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val
450 455 460

Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly
465 470 475 480

Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu
485 490 495

Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys
500 505 510

Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val
515 520 525

Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val
530 535 540

Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys
545 550 555 560

Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val
565 570 575

Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala
580 585 590

Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp

595 600 605

Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Cys

610 615 620

Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile

625 630 635 640

Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val

645 650 655

Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr

660 665 670

Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg

675 680 685

Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala

690 695 700

Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly

705 710 715 720

His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe

725 730 735

Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr

740 745 750

Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr

755 760 765

Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys

770 775 780

Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile

785 790 795 800

Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp

805 810 815

Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp

820 825 830

Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly

835 840 845

Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe
850 855 860

Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp
865 870 875 880

Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys
885 890 895

Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe
900 905 910

Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp
915 920 925

His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys
930 935 940

<210> 49

<211> 2823

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 49

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaagga tgcccacaag agcgaggtgg cccacagatt caaggatctg 180

ggcgaagaga acttcaaggc cctggtgctg atcgccttcg ctcagtatct ccagcagagc 240

cctttcgagg accacgtgaa gctggtcaac gaagtgaccg agttcgccaa gacctgtgtg 300

gccgatgaga gcgccgagaa ctgtgataag agcctgcaca ccctgttcgg cgacaagctg 360

tgtacagtgg ccacactgag agaaacctac ggcgagatgg ccgactgctg tgccaagcaa 420

gagcccgaga gaaacgagtg cttcctccag cacaaggacg acaaccccaa cctgcctaga 480

ctcgtgcgac ccgaagtgga tgtgatgtgc accgcctttc acgacaacga ggaaaccttc 540

ctgaagaagt acctgtacga gatcgccaga cggcacccct acttttatgc ccctgagctg 600

ctgttcttcg ccaagcggta taaggccgcc ttcaccgaat gttgccaggc cgctgataag 660

gctgcctgtc tgctgcctaa gctggacgag ctgagagatg agggcaaagc cagctctgcc 720

aagcagagac tgaaatgcgc cagcctccag aagttcggcg agagagcttt taaggcctgg 780

gccgttgcca gactgagcca gagatttcct aaggccgagt ttgccgaggt gtccaagctc 840

gtgaccgatc tgacaaaggt gcacaccgag tgctgtcacg gcgatctgct ggaatgtgcc 900

gacgatagag ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 960

ctgaaagagt gctgcgagaa gcccctgctg gaaaagtctc actgtatcgc cgaggtggaa 1020

aacgacgaga tgcctgccga tctgcctagc ctggctgccg atttcgtgga aagcaaggac 1080

gtgtgcaaga actacgccga ggccaaggat gtgtttctgg gcatgtttct gtatgagtac 1140

gcccgcagac accccgacta ttctgtggtt ctgctgctgc ggctggccaa aacctacgag 1200

acaaccctgg aaaaatgctg cgccgctgcc gatcctcacg agtgttatgc caaggtgttc 1260

gacgagttca agcctctggt ggaagaaccc cagaacctga tcaagcagaa ctgcgagctg 1320

ttcgagcagc tgggcgagta caagttccag aatgccctgc tcgtgcggta caccaagaaa 1380

gtgcctcagg tgtccacacc tacactggtt gaggtgtccc ggaatctggg caaagtgggc 1440

agcaagtgtt gcaagcaccc tgaggccaag agaatgcctt gcgccgagga ttacctgagc 1500

gtggtgctga atcagctgtg cgtgctgcac gagaaaaccc ctgtgtccga cagagtgacc 1560

aagtgctgta ccgagagcct cgtgaacaga aggccttgct ttagcgccct ggaagtggac 1620

gagacatacg tgcccaaaga gttcaacgcc gagacattca ccttccacgc cgacatctgc 1680

accctgtccg agaaagagcg gcagatcaag aagcagacag ccctggtcga gctggttaag 1740

cacaagccca aggccaccaa agaacagctg aaggccgtga tggacgactt cgccgccttt 1800

gtcgagaagt gctgcaaggc cgacgacaaa gagacatgct tcgccgaaga gggcaagaaa 1860

ctggtggcct gtgtggaacc cctcggcatg gaaaacggca atatcgccaa tagccagatt 1920

gccgccagca gcgtcagagt gacatttctg ggactgcaac actgggtgcc cgagctggct 1980

agactgaata gagccggcat ggtcaacgcc tggacaccca gcagcaacga cgataatccc 2040

tggattcaag tgaacctgct gcggcgtatg tgggtcacag gtgttgttac acagggcgca 2100

agcagactgg ccagccacga gtatctgaag gcctttaagg tggcctacag cctgaacggc 2160

cacgagttcg acttcatcca cgacgtgaac aagaagcaca aagagtttgt cggcaactgg 2220

aacaagaacg ccgtgcacgt gaacctgttc gagacacctg tggaagccca gtacgtgcgg 2280

ctgtacccta caagctgtca caccgcctgc actctgagat tcgaactgct gggatgcgag 2340

ctgaacggct gtgctaatcc tctgggcctg aagaacaaca gcatccccga taagcagatc 2400

accgccagct ccagctataa gacatggggc ctgcacctgt tcagctggaa cccttcttac 2460

gccagactgg acaagcaggg caacttcaat gcttgggtgg ccggcagcta cggcaatgat 2520

cagtggctgc aagtggacct gggcagcagc aaagaagtga caggcatcat cacccaaggg 2580

gccagaaatt tcggcagcgt gcagttcgtg gccagctaca aagtggccta ctccaacgac 2640

agcgccaact ggaccgagta tcaggaccct agaaccggca gctccaagat cttccccggc 2700

aattgggaca accacagcca caagaagaat ctgttcgaaa cccctatcct ggccagatat 2760

gtgcgcattc tgcccgtggc ctggcacaac agaattgccc tgagactgga actgctcggc 2820

tgc 2823

<210> 50

<211> 963

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 50

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
100 105 110

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
115 120 125

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro

130 135 140

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu

145 150 155 160

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His

165 170 175

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg

180 185 190

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg

195 200 205

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala

210 215 220

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser

225 230 235 240

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu

245 250 255

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro

260 265 270

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

275 280 285

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp

290 295 300

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser

305 310 315 320

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His

325 330 335

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

340 345 350

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

355 360 365

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg

370 375 380

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

625 630 635 640

Gly Ser Gly Gly Ser Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn

645 650 655

Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu

660 665 670

Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly

675 680 685

Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile

690 695 700

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

705 710 715 720

Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val

725 730 735

Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn

740 745 750

Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His

755 760 765

Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr

770 775 780

Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly

785 790 795 800

Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser

805 810 815

Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly

820 825 830

Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln

835 840 845

Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp

850 855 860

Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr

865 870 875 880

Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys
885 890 895
Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro
900 905 910
Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser
915 920 925
His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg
930 935 940
Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu
945 950 955 960
Leu Gly Cys

<210> 51
<211> 2889
<212> ADN
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"
<400> 51
ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60
gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120
aaccactgcg agacaaaagg atccgatgct cacaagagcg aggtggccca cagattcaag 180
gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240
cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300
tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360
aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420
aagcaagagc ccgagagaaa cgagtgcttc ctgcagcaca aggacgacaa ccccaacctg 480
cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540
accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgccccc 600
gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa gtgcgccagc ctgcagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgtaccc tgagcgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaggctgct ctcggagtgg gaggaagcgg aggatctggc 1920

ggttccggag gctcttgtgt ggaacccctc ggcatggaaa acggcaatat cgccaatagc 1980

cagattgccg ccagcagcgt cagagtgaca tttctgggac tgcagcactg ggtgcccgag 2040

ctggctagac tgaatagagc cggcatggtc aacgcctgga cacccagcag caacgacgat 2100

aacccttgga ttcaagtgaa cctgctgcgg cgtatgtggg tcacaggtgt tgttacacag 2160

ggcgcctcta gactggccag ccacgagtat ctgaaggcct ttaaggtggc ctacagcctg 2220

aacggccacg agttcgactt catccacgac gtgaacaaga agcacaaaga gtttgtcggc 2280

aactggaaca agaacgccgt gcacgtgaac ctgttcgaga cacctgtgga agcccagtac 2340

gtgcggctgt accctacaag ctgtcacacc gcctgcactc tgagattcga actgctggga 2400

tgcgagctga acggctgtgc taatcctctg ggcctgaaga acaacagcat ccccgataag 2460

cagatcaccg ccagctccag ctataagaca tggggcctgc acctgttcag ctggaaccct 2520

tcttacgcca gactggacaa gcagggcaac ttcaatgctt gggtggccgg cagctacggc 2580

aatgatcagt ggctgcaagt ggacctgggc agcagcaaag aagtgacagg catcatcacc 2640

cagggcgcca gaaatttcgg cagcgtgcag tttgtggcca gctacaaagt ggcctactcc 2700

aacgacagcg ccaactggac cgagtatcag gaccctagaa ccggcagctc caagatcttc 2760

cccggcaatt gggacaacca cagccacaag aagaatctgt tcgaaacccc tatcctggcc 2820

agatatgtgc gcattctgcc cgtggcctgg cacaacagaa ttgccctgag actggaactg 2880

ctcggctgt 2889

<210> 52

<211> 961

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 52

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
100 105 110

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
115 120 125

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

370 375 380

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr

385 390 395 400

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu

405 410 415

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

420 425 430

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu

435 440 445

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro

450 455 460

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys

465 470 475 480

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys

485 490 495

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His

500 505 510

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser

515 520 525

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr

530 535 540

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp

545 550 555 560

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala

565 570 575

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu

580 585 590

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys

595 600 605

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val

610 615 620

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

865 870 875 880

Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala

885 890 895

Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr

900 905 910

Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys

915 920 925

Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu

930 935 940

Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly

945 950 955 960

Cys

<210> 53

<211> 2883

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 53

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccgatgct cacaagagcg aggtggccca cagattcaag 180

gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240

cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300

tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360

aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420

aagcaagagc ccgagagaaa cgagtgcttc ctgcagcaca aggacgacaa ccccaacctg 480

cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540

accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgccccc 600

gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa gtgcgccagc ctgcagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgtaccc tgagcgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaggctgct ctcggaggaa gcggaggatc tggcggttcc 1920

ggaggctctt gtgtggaacc cctcggcatg gaaaacggca atatcgccaa tagccagatt 1980

gccgccagca gcgtcagagt gacatttctg ggactgcagc actgggtgcc cgagctggct 2040

agactgaata gagccggcat ggtcaacgcc tggacaccca gcagcaacga cgataaccct 2100

tggattcaag tgaacctgct gcggcgtatg tgggtcacag gtgttgttac acagggcgcc 2160

tctagactgg ccagccacga gtatctgaag gcctttaagg tggcctacag cctgaacggc 2220

cacgagttcg acttcatcca cgacgtgaac aagaagcaca aagagtttgt cggcaactgg 2280

aacaagaacg ccgtgcacgt gaacctgttc gagacacctg tggaagccca gtacgtgcgg 2340

ctgtacccta caagctgtca caccgcctgc actctgagat tcgaactgct gggatgcgag 2400

ctgaacggct gtgctaatcc tctgggcctg aagaacaaca gcatccccga taagcagatc 2460

accgccagct ccagctataa gacatggggc ctgcacctgt tcagctggaa cccttcttac 2520

gccagactgg acaagcaggg caacttcaat gcttgggtgg ccggcagcta cggcaatgat 2580

cagtggctgc aagtggacct gggcagcagc aaagaagtga caggcatcat cacccagggc 2640

gccagaaatt tcggcagcgt gcagtttgtg gccagctaca aagtggccta ctccaacgac 2700

agcgccaact ggaccgagta tcaggaccct agaaccggca gctccaagat cttccccggc 2760

aattgggaca accacagcca caagaagaat ctgttcgaaa cccctatcct ggccagatat 2820

gtgcgcattc tgcccgtggc ctggcacaac agaattgccc tgagactgga actgctcggc 2880

tgt 2883

<210> 54

<211> 954

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 54

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
100 105 110

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu

115 120 125

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro

130 135 140

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu

145 150 155 160

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His

165 170 175

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg

180 185 190

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg

195 200 205

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala

210 215 220

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser

225 230 235 240

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu

245 250 255

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro

260 265 270

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

275 280 285

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp

290 295 300

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser

305 310 315 320

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His

325 330 335

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

340 345 350

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

355 360 365

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
370 375 380
Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
385 390 395 400
Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
405 410 415
Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
420 425 430
Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
435 440 445
Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
450 455 460
Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
465 470 475 480
Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
485 490 495
Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
500 505 510
Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
515 520 525
Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
530 535 540
Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
545 550 555 560
Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
565 570 575
Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
580 585 590
Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
595 600 605
Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val

610 615 620

Ala Ala Ser Gln Ala Ala Leu Gly Gly Gly Gly Ser Cys Val Glu Pro

625 630 635 640

Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile Ala Ala Ser

645 650 655

Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val Pro Glu Leu

660 665 670

Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr Pro Ser Ser

675 680 685

Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg Arg Met Trp

690 695 700

Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala Ser His Glu

705 710 715 720

Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly His Glu Phe

725 730 735

Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe Val Gly Asn

740 745 750

Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr Pro Val Glu

755 760 765

Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr Ala Cys Thr

770 775 780

Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys Ala Asn Pro

785 790 795 800

Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile Thr Ala Ser

805 810 815

Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp Asn Pro Ser

820 825 830

Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp Val Ala Gly

835 840 845

Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly Ser Ser Lys

850 855 860

Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly Ser Val
865 870 875 880

Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser Ala Asn
885 890 895

Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys Ile Phe Pro
900 905 910

Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe Glu Thr Pro
915 920 925

Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp His Asn Arg
930 935 940

Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys
945 950

<210> 55

<211> 2862

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 55

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atccgatgct cacaagagcg aggtggccca cagattcaag 180

gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240

cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300

tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360

aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420

aagcaagagc ccgagagaaa cgagtgcttc ctgcagcaca aggacgacaa ccccaacctg 480

cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540

accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgcccct 600

gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa gtgcgccagc ctgcagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgtaccc tgagcgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaagctgct ctcggaggcg gaggatcttg tgtggaaccc 1920

ctcggcatgg aaaacggcaa tatcgccaat agccagattg ccgccagcag cgtcagagtg 1980

acatttctgg gactgcagca ctgggtgccc gagctggcta gactgaatag agccggcatg 2040

gtcaacgcct ggacacccag cagcaacgac gataaccctt ggattcaagt gaacctgctg 2100

cggcgtatgt gggtcacagg tgttgttaca cagggcgcct ctagactggc cagccacgag 2160

tatctgaagg cctttaaggt ggcctacagc ctgaacggcc acgagttcga cttcatccac 2220

gacgtgaaca agaagcacaa agagtttgtc ggcaactgga acaagaacgc cgtgcacgtg 2280

aacctgttcg agacacctgt ggaagcccag tacgtgcggc tgtaccctac aagctgtcac 2340

accgcctgca ctctgagatt cgaactgctg ggatgcgagc tgaacggctg tgctaatcct 2400

ctgggcctga agaacaacag catccccgat aagcagatca ccgccagctc cagctataag 2460

acatggggcc tgcacctgtt cagctggaac ccttcttacg ccagactgga caagcagggc 2520

aacttcaatg cttgggtggc cggcagctac ggcaatgatc agtggctgca agtggacctg 2580

ggcagcagca aagaagtgac aggcatcatc acccagggcg ccagaaattt cggcagcgtg 2640

cagtttgtgg ccagctacaa agtggcctac tccaacgaca gcgccaactg gaccgagtat 2700

caggacccta gaaccggcag ctccaagatc ttccccggca attgggacaa ccacagccac 2760

aagaagaatc tgttcgaaac ccctatcctg gccagatatg tgcgcattct gcccgtggcc 2820

tggcacaaca gaattgccct gagactggaa ctgctcggct gc 2862

<210> 56

<211> 959

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 56

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
100 105 110

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
115 120 125

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro

130 135 140

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu

145 150 155 160

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His

165 170 175

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg

180 185 190

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg

195 200 205

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala

210 215 220

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser

225 230 235 240

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu

245 250 255

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro

260 265 270

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

275 280 285

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp

290 295 300

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser

305 310 315 320

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His

325 330 335

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

340 345 350

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

355 360 365

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg

370 375 380

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

625 630 635 640

Ser Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser

645 650 655

Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His

660 665 670

Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala

675 680 685

Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu

690 695 700

Leu Arg Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg

705 710 715 720

Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu

725 730 735

Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys

740 745 750

Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe

755 760 765

Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys

770 775 780

His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn

785 790 795 800

Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys

805 810 815

Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe

820 825 830

Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn

835 840 845

Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp

850 855 860

Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg

865 870 875 880

Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser
885 890 895
Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser
900 905 910
Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn
915 920 925
Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val
930 935 940
Ala Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys
945 950 955

<210> 57
<211> 2877
<212> ADN
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 57
ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60
gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120
aaccactgcg agacaaaagg atccgatgct cacaagagcg aggtggccca cagattcaag 180
gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240
cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300
tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360
aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420
aagcaagagc ccgagagaaa cgagtgcttc ctgcagcaca aggacgacaa ccccaacctg 480
cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540
accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgccccc 600
gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660
gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720
tctgccaagc agagactgaa gtgcgccagc ctgcagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgtaccc tgagcgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaagctgct ctcggaggcg gaggctccgg aggcggagga 1920

tcttgtgtgg aacccctcgg catggaaaac ggcaatatcg ccaatagcca gattgccgcc 1980

agcagcgtca gagtgacatt tctgggactg cagcactggg tgcccgagct ggctagactg 2040

aatagagccg gcatggtcaa cgcctggaca cccagcagca acgacgataa cccttggatt 2100

caagtgaacc tgctgcggcg tatgtgggtc acaggtgttg ttacacaggg cgcctctaga 2160

ctggccagcc acgagtatct gaaggccttt aaggtggcct acagcctgaa cggccacgag 2220

ttcgacttca tccacgacgt gaacaagaag cacaaagagt ttgtcggcaa ctggaacaag 2280

aacgccgtgc acgtgaacct gttcgagaca cctgtggaag cccagtacgt gcggctgtac 2340

cctacaagct gtcacaccgc ctgcactctg agattcgaac tgctgggatg cgagctgaac 2400

ggctgtgcta atcctctggg cctgaagaac aacagcatcc ccgataagca gatcaccgcc 2460

agctccagct ataagacatg gggcctgcac ctgttcagct ggaacccttc ttacgccaga 2520

ctggacaagc agggcaactt caatgcttgg gtggccggca gctacggcaa tgatcagtgg 2580

ctgcaagtgg acctgggcag cagcaaagaa gtgacaggca tcatcaccca gggcgccaga 2640

aatttcggca gcgtgcagtt tgtggccagc tacaaagtgg cctactccaa cgacagcgcc 2700

aactggaccg agtatcagga ccctagaacc ggcagctcca agatcttccc cggcaattgg 2760

gacaaccaca gccacaagaa gaatctgttc gaaaccccta tcctggccag atatgtgcgc 2820

attctgcccg tggcctggca caacagaatt gccctgagac tggaactgct cggctgc 2877

<210> 58

<211> 949

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 58

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala
35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn
50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser
65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala
85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu
100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu
115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg
130 135 140

Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg

145 150 155 160

Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn

165 170 175

Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His

180 185 190

Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys

195 200 205

Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu

210 215 220

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

225 230 235 240

Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala

245 250 255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala

260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His

275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala

290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys

305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile

325 330 335

Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala

340 345 350

Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala

355 360 365

Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His

370 375 380

Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu

385 390 395 400

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

645 650 655

Phe Leu Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg

660 665 670

Ala Gly Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro

675 680 685

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

690 695 700

Thr Gln Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe

705 710 715 720

Lys Val Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp

725 730 735

Val Asn Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala

740 745 750

Val His Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg

755 760 765

Leu Tyr Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu

770 775 780

Leu Gly Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn

785 790 795 800

Asn Ser Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr

805 810 815

Trp Gly Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp

820 825 830

Lys Gln Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp

835 840 845

Gln Trp Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile

850 855 860

Ile Thr Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser

865 870 875 880

Tyr Lys Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln

885 890 895

Asp Pro Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn
900 905 910

His Ser His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr
915 920 925

Val Arg Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg Leu
930 935 940

Glu Leu Leu Gly Cys
945

<210> 59

<211> 2847

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 59

ctggacatct gcagcaagaa tccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaagga tgcccacaag agcgaggtgg cccacagatt caaggatctg 180

ggcgaagaga acttcaaggc cctggtgctg atcgccttcg ctcagtatct ccagcagagc 240

cctttcgagg accacgtgaa gctggtcaac gaagtgaccg agttcgccaa gacctgtgtg 300

gccgatgaga gcgccgagaa ctgtgataag agcctgcaca ccctgttcgg cgacaagctg 360

tgtacagtgg ccacactgag agaaacctac ggcgagatgg ccgactgctg tgccaagcaa 420

gagcccgaga gaaacgagtg cttcctgcag cacaaggacg acaaccccaa cctgcctaga 480

ctcgtgcgac ccgaagtgga tgtgatgtgc accgcctttc acgacaacga ggaaaccttc 540

ctgaagaagt acctgtacga gatcgccaga cggcaccct actttatgc ccctgagctg 600

ctgttcttcg ccaagcggta taaggccgcc ttcaccgaat gttgccaggc cgctgataag 660

gctgcctgtc tgctgcctaa gctggacgag ctgagagatg agggcaaagc cagctctgcc 720

aagcagagac tgaagtgcgc cagcctgcag aagttcggcg agagagcttt taaggcctgg 780

gccgttgcca gactgagcca gagatttcct aaggccgagt ttgccgaggt gtccaagctc 840

gtgaccgatc tgacaaaggt gcacaccgag tgctgtcacg gcgatctgct ggaatgtgcc 900

gacgatagag ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 960

ctgaaagagt gctgcgagaa gcccctgctg gaaaagtctc actgtatcgc cgaggtggaa 1020

aacgacgaga tgcctgccga tctgcctagc ctggctgccg atttcgtgga aagcaaggac 1080

gtgtgcaaga actacgccga ggccaaggat gtgtttctgg gcatgtttct gtatgagtac 1140

gcccgcagac accccgacta ttctgtggtt ctgctgctgc ggctggccaa aacctacgag 1200

acaaccctgg aaaaatgctg cgccgctgcc gatcctcacg agtgttatgc caaggtgttc 1260

gacgagttca agcctctggt ggaagaaccc cagaacctga tcaagcagaa ctgcgagctg 1320

ttcgagcagc tgggcgagta caagttccag aatgccctgc tcgtgcggta caccaagaaa 1380

gtgcctcagg tgtccacacc tacactggtt gaggtgtccc ggaatctggg caaagtgggc 1440

agcaagtgtt gcaagcaccc tgaggccaag agaatgcctt gcgccgagga ttacctgagc 1500

gtggtgctga atcagctgtg cgtgctgcac gagaaaaccc ctgtgtccga cagagtgacc 1560

aagtgctgta ccgagagcct cgtgaacaga aggccttgct ttagcgccct ggaagtggac 1620

gagacatacg tgcccaaaga gttcaacgcc gagacattca ccttccacgc cgacatctgt 1680

accctgagcg agaaagagcg gcagatcaag aagcagacag ccctggtcga gctggttaag 1740

cacaagccca aggccaccaa agaacagctg aaggccgtga tggacgactt cgccgccttt 1800

gtcgagaagt gctgcaaggc cgacgacaaa gagacatgct tcgccgaaga gggcaagaaa 1860

ctggtggctg cttctcaggc tgccctggga ctgtgtgtgg aacccctcgg catggaaaac 1920

ggcaatatcg ccaatagcca gattgccgcc agcagcgtca gagtgacatt tctgggactg 1980

cagcactggg tgcccgagct ggctagactg aatagagccg gcatggtcaa cgcctggaca 2040

cccagcagca acgacgataa tccctggatc caagtgaacc tgctgcggcg tatgtgggtc 2100

acaggtgttg ttacacaggg cgcctctaga ctggccagcc acgagtatct gaaggccttt 2160

aaggtggcct acagcctgaa cggccacgag ttcgacttca tccacgacgt gaacaagaag 2220

cacaaagagt ttgtcggcaa ctggaacaag aacgccgtgc acgtgaacct gttcgagaca 2280

cctgtggaag cccagtacgt gcggctgtac cctacaagct gtcacaccgc ctgcactctg 2340

agattcgaac tgctgggatg cgagctgaac ggctgtgcta atcctctggg cctgaagaac 2400

aacagcatcc ccgataagca gatcaccgcc agctccagct ataagacatg gggcctgcac 2460

ctgttcagct ggaacccttc ttacgccaga ctggacaagc agggcaactt caatgcttgg 2520

gtggccggca gctacggcaa tgatcagtgg ctgcaagtgg acctgggcag cagcaaagaa 2580

gtgacaggca tcatcaccca gggcgccaga aatttcggca gcgtgcagtt tgtggccagc 2640

tacaaagtgg cctactccaa cgacagcgcc aactggaccg agtaccagga tcctagaacc 2700

ggcagctcca agatcttccc cggcaattgg gacaaccaca gccacaagaa gaatctgttc 2760

gaaaccccta tcctggccag atatgtgcgg attctgcccg tggcctggca caacagaatt 2820

gccctgagac tggaactgct cggctgt 2847

<210> 60

<400> 60

000

<210> 61

<400> 61

000

<210> 62

<211> 12

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Peptit tổng hợp"

<400> 62

Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser

1 5 10

<210> 63

<211> 8

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Peptit tổng hợp"

<400> 63

Gly Ser Gly Ser Gly Ser Gly Ser

1 5

<210> 64

<211> 5

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Peptit tổng hợp"

<400> 64

Gly Gly Gly Gly Ser

1 5

<210> 65

<211> 10

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Peptit tổng hợp"

<400> 65

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10

<210> 66

<211> 8

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Peptit tổng hợp"

<400> 66

Gly Ser His His His His His His

1 5

<210> 67

<211> 6

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Đuôi 6xHis tổng hợp"

<400> 67

His His His His His His

1 5

<210> 68

<211> 463

<212> PRT

<213> Chuột nhắt nhà

<400> 68

Met Gln Val Ser Arg Val Leu Ala Ala Leu Cys Gly Met Leu Leu Cys

1 5 10 15

Ala Ser Gly Leu Phe Ala Ala Ser Gly Asp Phe Cys Asp Ser Ser Leu

20 25 30

Cys Leu Asn Gly Gly Thr Cys Leu Thr Gly Gln Asp Asn Asp Ile Tyr

35 40 45

Cys Leu Cys Pro Glu Gly Phe Thr Gly Leu Val Cys Asn Glu Thr Glu

50 55 60

Arg Gly Pro Cys Ser Pro Asn Pro Cys Tyr Asn Asp Ala Lys Cys Leu

65 70 75 80

Val Thr Leu Asp Thr Gln Arg Gly Asp Ile Phe Thr Glu Tyr Ile Cys

85 90 95

Gln Cys Pro Val Gly Tyr Ser Gly Ile His Cys Glu Thr Glu Thr Asn

100 105 110

Tyr Tyr Asn Leu Asp Gly Glu Tyr Met Phe Thr Thr Ala Val Pro Asn

115 120 125

Thr Ala Val Pro Thr Pro Ala Pro Thr Pro Asp Leu Ser Asn Asn Leu

130 135 140

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

145 150 155 160

Asp Ser Gln Ile Ser Ala Ser Ser Val Tyr Met Gly Phe Met Gly Leu
165 170 175

Gln Arg Trp Gly Pro Glu Leu Ala Arg Leu Tyr Arg Thr Gly Ile Val
180 185 190

Asn Ala Trp Thr Ala Ser Asn Tyr Asp Ser Lys Pro Trp Ile Gln Val
195 200 205

Asn Leu Leu Arg Lys Met Arg Val Ser Gly Val Met Thr Gln Gly Ala
210 215 220

Ser Arg Ala Gly Arg Ala Glu Tyr Leu Lys Thr Phe Lys Val Ala Tyr
225 230 235 240

Ser Leu Asp Gly Arg Lys Phe Glu Phe Ile Gln Asp Glu Ser Gly Gly
245 250 255

Asp Lys Glu Phe Leu Gly Asn Leu Asp Asn Asn Ser Leu Lys Val Asn
260 265 270

Met Phe Asn Pro Thr Leu Glu Ala Gln Tyr Ile Lys Leu Tyr Pro Val
275 280 285

Ser Cys His Arg Gly Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu
290 295 300

Leu His Gly Cys Ser Glu Pro Leu Gly Leu Lys Asn Asn Thr Ile Pro
305 310 315 320

Asp Ser Gln Met Ser Ala Ser Ser Ser Tyr Lys Thr Trp Asn Leu Arg
325 330 335

Ala Phe Gly Trp Tyr Pro His Leu Gly Arg Leu Asp Asn Gln Gly Lys
340 345 350

Ile Asn Ala Trp Thr Ala Gln Ser Asn Ser Ala Lys Glu Trp Leu Gln
355 360 365

Val Asp Leu Gly Thr Gln Arg Gln Val Thr Gly Ile Ile Thr Gln Gly
370 375 380

Ala Arg Asp Phe Gly His Ile Gln Tyr Val Ala Ser Tyr Lys Val Ala
385 390 395 400

His Ser Asp Asp Gly Val Gln Trp Thr Val Tyr Glu Glu Gln Gly Ser

405 410 415

Ser Lys Val Phe Gln Gly Asn Leu Asp Asn Asn Ser His Lys Lys Asn

420 425 430

Ile Phe Glu Lys Pro Phe Met Ala Arg Tyr Val Arg Val Leu Pro Val

435 440 445

Ser Trp His Asn Arg Ile Thr Leu Arg Leu Glu Leu Leu Gly Cys

450 455 460

<210> 69

<211> 1030

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 69

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro

1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr

20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu

35 40 45

Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly

50 55 60

Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr

65 70 75 80

Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn

85 90 95

Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr

100 105 110

Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly

115 120 125

Arg Asn Cys Gln Tyr Lys Asp Ala His Lys Ser Glu Val Ala His Arg
130 135 140

Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala
145 150 155 160

Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu
165 170 175

Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser
180 185 190

Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu
195 200 205

Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys
210 215 220

Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys
225 230 235 240

Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val
245 250 255

Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr
260 265 270

Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu
275 280 285

Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln
290 295 300

Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg
305 310 315 320

Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser
325 330 335

Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg
340 345 350

Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu
355 360 365

Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu

370 375 380

Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu

385 390 395 400

Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro

405 410 415

Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met

420 425 430

Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp

435 440 445

Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe

450 455 460

Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu

465 470 475 480

Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala

485 490 495

Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys

500 505 510

Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu

515 520 525

Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg

530 535 540

Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val

545 550 555 560

Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu

565 570 575

Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn

580 585 590

Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr

595 600 605

Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala

610 615 620

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

865 870 875 880

Ser Gly His Ile Gln Asp Tyr Gln Ile Thr Ala Ser Ser Ile Phe Arg

885 890 895

Thr Leu Asn Met Asp Met Phe Thr Trp Glu Pro Arg Lys Ala Arg Leu

900 905 910

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

915 920 925

Ser Gln Trp Leu Gln Val Asp Leu Leu Val Pro Thr Lys Val Thr Gly

930 935 940

Ile Ile Thr Gln Gly Ala Lys Asp Phe Gly His Val Gln Phe Val Gly

945 950 955 960

Ser Tyr Lys Leu Ala Tyr Ser Asn Asp Gly Glu His Trp Thr Val Tyr

965 970 975

Gln Asp Glu Lys Gln Arg Lys Asp Lys Val Phe Gln Gly Asn Phe Asp

980 985 990

Asn Asp Thr His Arg Lys Asn Val Ile Asp Pro Pro Ile Tyr Ala Arg

995 1000 1005

His Ile Arg Ile Leu Pro Trp Ser Trp Tyr Gly Arg Ile Thr Leu

1010 1015 1020

Arg Ser Glu Leu Leu Gly Cys

1025 1030

<210> 70

<211> 1036

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 70

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro

1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30
Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45
Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly
50 55 60
Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr
65 70 75 80
Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn
85 90 95
Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr
100 105 110
Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly
115 120 125
Arg Asn Cys Gln Tyr Lys Asp Ala His Lys Ser Glu Val Ala His Arg
130 135 140
Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala
145 150 155 160
Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu
165 170 175
Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser
180 185 190
Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu
195 200 205
Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys
210 215 220
Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys
225 230 235 240
Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val
245 250 255
Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr

260 265 270

Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu

275 280 285

Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln

290 295 300

Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg

305 310 315 320

Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser

325 330 335

Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg

340 345 350

Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu

355 360 365

Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu

370 375 380

Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu

385 390 395 400

Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro

405 410 415

Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met

420 425 430

Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp

435 440 445

Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe

450 455 460

Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu

465 470 475 480

Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala

485 490 495

Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys

500 505 510

Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu
515 520 525

Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg
530 535 540

Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val
545 550 555 560

Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu
565 570 575

Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn
580 585 590

Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr
595 600 605

Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala
610 615 620

Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr
625 630 635 640

Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln
645 650 655

Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys
660 665 670

Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe
675 680 685

Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu
690 695 700

Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Cys Ser Gly
705 710 715 720

Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn Gln Gln Ile Thr Ala
725 730 735

Ser Ser Thr His Arg Ala Leu Phe Gly Leu Gln Lys Trp Tyr Pro Tyr
740 745 750

Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp Thr Ala Ala

755 760 765

Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln Arg Lys Met

770 775 780

Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile Gly Ser Pro

785 790 795 800

Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp Gly Lys Thr

805 810 815

Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Asp Met Val Phe Arg

820 825 830

Gly Asn Ile Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe Thr Pro Pro

835 840 845

Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Val Cys Arg Arg His

850 855 860

Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser Gly Cys Ser

865 870 875 880

Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr Gln Ile Thr

885 890 895

Ala Ser Ser Ile Phe Arg Thr Leu Asn Met Asp Met Phe Thr Trp Glu

900 905 910

Pro Arg Lys Ala Arg Leu Asp Lys Gln Gly Lys Val Asn Ala Trp Thr

915 920 925

Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp Leu Leu Val

930 935 940

Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys Asp Phe Gly

945 950 955 960

His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser Asn Asp Gly

965 970 975

Glu His Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys Asp Lys Val

980 985 990

Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn Val Ile Asp

995 1000 1005

Pro Pro Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp Ser Trp
1010 1015 1020
Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys
1025 1030 1035

<210> 71
<211> 872
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 71

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

145 150 155 160

Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu

165 170 175

Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser

180 185 190

Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu

195 200 205

Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys

210 215 220

Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys

225 230 235 240

Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val

245 250 255

Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr

260 265 270

Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu

275 280 285

Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln

290 295 300

Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg

305 310 315 320

Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser

325 330 335

Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg

340 345 350

Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu

355 360 365

Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu

370 375 380

Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu

385 390 395 400

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

645 650 655

Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys

660 665 670

Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe

675 680 685

Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu

690 695 700

Glu Gly Lys Lys Leu Val Ala Cys Ser Gly Pro Leu Gly Ile Glu Gly

705 710 715 720

Gly Ile Ile Ser Asn Gln Gln Ile Thr Ala Ser Ser Thr His Arg Ala

725 730 735

Leu Phe Gly Leu Gln Lys Trp Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys

740 745 750

Lys Gly Leu Ile Asn Ala Trp Thr Ala Ala Glu Asn Asp Arg Trp Pro

755 760 765

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

770 775 780

Thr Gln Gly Ala Lys Arg Ile Gly Ser Pro Glu Tyr Ile Lys Ser Tyr

785 790 795 800

Lys Ile Ala Tyr Ser Asn Asp Gly Lys Thr Trp Ala Met Tyr Lys Val

805 810 815

Lys Gly Thr Asn Glu Asp Met Val Phe Arg Gly Asn Ile Asp Asn Asn

820 825 830

Thr Pro Tyr Ala Asn Ser Phe Thr Pro Pro Ile Lys Ala Gln Tyr Val

835 840 845

Arg Leu Tyr Pro Gln Val Cys Arg Arg His Cys Thr Leu Arg Met Glu

850 855 860

Leu Leu Gly Cys Glu Leu Ser Gly

865 870

<210> 72

<211> 878

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 72

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45

Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly
50 55 60

Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr
65 70 75 80

Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn
85 90 95

Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr
100 105 110

Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly
115 120 125

Arg Asn Cys Gln Tyr Lys Asp Ala His Lys Ser Glu Val Ala His Arg
130 135 140

Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala
145 150 155 160

Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu
165 170 175

Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser
180 185 190

Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu

195 200 205

Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys

210 215 220

Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys

225 230 235 240

Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val

245 250 255

Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr

260 265 270

Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu

275 280 285

Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln

290 295 300

Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg

305 310 315 320

Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser

325 330 335

Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg

340 345 350

Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu

355 360 365

Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu

370 375 380

Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu

385 390 395 400

Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro

405 410 415

Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met

420 425 430

Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp

435 440 445

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

690 695 700

Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Cys Ser Gly

705 710 715 720

Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn Gln Gln Ile Thr Ala

725 730 735

Ser Ser Thr His Arg Ala Leu Phe Gly Leu Gln Lys Trp Tyr Pro Tyr

740 745 750

Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp Thr Ala Ala

755 760 765

Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln Arg Lys Met

770 775 780

Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile Gly Ser Pro

785 790 795 800

Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp Gly Lys Thr

805 810 815

Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Asp Met Val Phe Arg

820 825 830

Gly Asn Ile Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe Thr Pro Pro

835 840 845

Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Val Cys Arg Arg His

850 855 860

Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser Gly

865 870 875

<210> 73

<211> 783

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 73

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15
Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30
Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala
35 40 45
His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn
50 55 60
Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser
65 70 75 80
Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala
85 90 95
Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu
100 105 110
His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu
115 120 125
Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg
130 135 140
Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg
145 150 155 160
Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn
165 170 175
Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His
180 185 190
Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys
195 200 205
Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu
210 215 220
Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala
225 230 235 240
Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala

245 250 255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala

260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His

275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala

290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys

305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile

325 330 335

Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala

340 345 350

Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala

355 360 365

Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His

370 375 380

Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu

385 390 395 400

Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr

405 410 415

Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn

420 425 430

Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys

435 440 445

Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val

450 455 460

Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly

465 470 475 480

Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu

485 490 495

Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys
500 505 510
Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val
515 520 525
Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val
530 535 540
Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys
545 550 555 560
Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val
565 570 575
Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala
580 585 590
Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp
595 600 605
Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Cys
610 615 620
Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile
625 630 635 640
Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val
645 650 655
Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr
660 665 670
Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg
675 680 685
Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala
690 695 700
Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly
705 710 715 720
His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe
725 730 735
Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr

740 745 750

Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr

755 760 765

Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly

770 775 780

<210> 74

<211> 789

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 74

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu

1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys

20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala

35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn

50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser

65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala

85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu

100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu

115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg

130 135 140

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

385 390 395 400

Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr

405 410 415

Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn

420 425 430

Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys

435 440 445

Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val

450 455 460

Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly

465 470 475 480

Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu

485 490 495

Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys

500 505 510

Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val

515 520 525

Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val

530 535 540

Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys

545 550 555 560

Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val

565 570 575

Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala

580 585 590

Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp

595 600 605

Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala

610 615 620

Ser Gln Ala Ala Leu Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn

625 630 635 640

Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu
645 650 655
Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly
660 665 670
Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile
675 680 685
Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr Gly Val Val Thr Gln
690 695 700
Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val
705 710 715 720
Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn
725 730 735
Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His
740 745 750
Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr
755 760 765
Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly
770 775 780
Cys Glu Leu Asn Gly
785

<210> 75
<211> 387
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 75
Met Pro Arg Pro Arg Leu Leu Ala Ala Leu Cys Gly Ala Leu Leu Cys
1 5 10 15
Ala Pro Ser Leu Leu Val Ala Leu Asp Ile Cys Ser Lys Asn Pro Cys

20 25 30

His Asn Gly Gly Leu Cys Glu Glu Ile Ser Gln Glu Val Arg Gly Asp

35 40 45

Val Phe Pro Ser Tyr Thr Cys Thr Cys Leu Lys Gly Tyr Ala Gly Asn

50 55 60

His Cys Glu Thr Lys Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn

65 70 75 80

Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu

85 90 95

Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly

100 105 110

Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile

115 120 125

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

130 135 140

Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val

145 150 155 160

Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn

165 170 175

Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His

180 185 190

Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr

195 200 205

Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly

210 215 220

Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser

225 230 235 240

Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly

245 250 255

Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln

260 265 270

Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp
275 280 285
Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr
290 295 300
Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys
305 310 315 320
Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro
325 330 335
Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser
340 345 350
His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg
355 360 365
Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu
370 375 380
Leu Gly Cys
385

<210> 76
<211> 310
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 76
Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln
1 5 10 15
Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp
20 25 30
Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp
35 40 45
Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu

50 55 60

Arg Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu

65 70 75 80

Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn

85 90 95

Gly His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu

100 105 110

Phe Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu

115 120 125

Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His

130 135 140

Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly

145 150 155 160

Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln

165 170 175

Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser

180 185 190

Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala

195 200 205

Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu

210 215 220

Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn

225 230 235 240

Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn

245 250 255

Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser

260 265 270

Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu

275 280 285

Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala

290 295 300

Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys
305 310 315

<210> 77
<211> 134
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 77

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15
Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30
Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45
Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly
50 55 60
Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr
65 70 75 80
Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn
85 90 95
Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr
100 105 110
Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly
115 120 125
Arg Asn Cys Gln Tyr Lys
130

<210> 78
<211> 325
<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 78

Cys Ser Gly Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn Gln Gln
1 5 10 15

Ile Thr Ala Ser Ser Thr His Arg Ala Leu Phe Gly Leu Gln Lys Trp
20 25 30

Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp
35 40 45

Thr Ala Ala Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln
50 55 60

Arg Lys Met Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile
65 70 75 80

Gly Ser Pro Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp
85 90 95

Gly Lys Thr Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Asp Met
100 105 110

Val Phe Arg Gly Asn Ile Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe
115 120 125

Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Val Cys
130 135 140

Arg Arg His Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser
145 150 155 160

Gly Cys Ser Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr
165 170 175

Gln Ile Thr Ala Ser Ser Ile Phe Arg Thr Leu Asn Met Asp Met Phe
180 185 190

Thr Trp Glu Pro Arg Lys Ala Arg Leu Asp Lys Gln Gly Lys Val Asn
195 200 205

Ala Trp Thr Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp
210 215 220
Leu Leu Val Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys
225 230 235 240
Asp Phe Gly His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser
245 250 255
Asn Asp Gly Glu His Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys
260 265 270
Asp Lys Val Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn
275 280 285
Val Ile Asp Pro Pro Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp
290 295 300
Ser Trp Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys Thr
305 310 315 320
Glu Glu Glu

<210> 79
<211> 319
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 79
Cys Ser Gly Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn Gln Gln
1 5 10 15
Ile Thr Ala Ser Ser Thr His Arg Ala Leu Phe Gly Leu Gln Lys Trp
20 25 30
Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp
35 40 45
Thr Ala Ala Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln

50 55 60

Arg Lys Met Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile

65 70 75 80

Gly Ser Pro Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp

85 90 95

Gly Lys Thr Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Asp Met

100 105 110

Val Phe Arg Gly Asn Ile Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe

115 120 125

Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Val Cys

130 135 140

Arg Arg His Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser

145 150 155 160

Gly Cys Ser Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr

165 170 175

Gln Ile Thr Ala Ser Ser Ile Phe Arg Thr Leu Asn Met Asp Met Phe

180 185 190

Thr Trp Glu Pro Arg Lys Ala Arg Leu Asp Lys Gln Gly Lys Val Asn

195 200 205

Ala Trp Thr Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp

210 215 220

Leu Leu Val Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys

225 230 235 240

Asp Phe Gly His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser

245 250 255

Asn Asp Gly Gln His Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys

260 265 270

Asp Lys Val Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn

275 280 285

Val Ile Asp Pro Pro Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp

290 295 300

Ser Trp Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys
305 310 315

<210> 80

<211> 1029

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 80

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45

Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly
50 55 60

Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr
65 70 75 80

Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn
85 90 95

Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr
100 105 110

Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly
115 120 125

Arg Asn Cys Gln Tyr Lys Asp Ala His Lys Ser Glu Val Ala His Arg
130 135 140

Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala
145 150 155 160

Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu

165 170 175

Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser

180 185 190

Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu

195 200 205

Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys

210 215 220

Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys

225 230 235 240

Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val

245 250 255

Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr

260 265 270

Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu

275 280 285

Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln

290 295 300

Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg

305 310 315 320

Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser

325 330 335

Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg

340 345 350

Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu

355 360 365

Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu

370 375 380

Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu

385 390 395 400

Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro

405 410 415

Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met
420 425 430

Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp
435 440 445

Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe
450 455 460

Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu
465 470 475 480

Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala
485 490 495

Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys
500 505 510

Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu
515 520 525

Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg
530 535 540

Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val
545 550 555 560

Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu
565 570 575

Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn
580 585 590

Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr
595 600 605

Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala
610 615 620

Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr
625 630 635 640

Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln
645 650 655

Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys

660 665 670

Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe

675 680 685

Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu

690 695 700

Glu Gly Lys Lys Leu Val Ala Cys Val Glu Pro Leu Gly Met Glu Asn

705 710 715 720

Gly Asn Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr

725 730 735

Phe Leu Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg

740 745 750

Ala Gly Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro

755 760 765

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

770 775 780

Thr Gln Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe

785 790 795 800

Lys Val Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp

805 810 815

Val Asn Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala

820 825 830

Val His Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg

835 840 845

Leu Tyr Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu

850 855 860

Leu Gly Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn

865 870 875 880

Asn Ser Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr

885 890 895

Trp Gly Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp

900 905 910

Lys Gln Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp
915 920 925

Gln Trp Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile
930 935 940

Ile Thr Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser
945 950 955 960

Tyr Lys Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln
965 970 975

Asp Pro Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn
980 985 990

His Ser His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr
995 1000 1005

Val Arg Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg
1010 1015 1020

Leu Glu Leu Leu Gly Cys
1025

<210> 81

<211> 3087

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 81

gacatctgcg accccaatcc ttgcgagaat ggcggcattt gtctgcctgg actggccgat 60

ggcagcttct cttgtgaatg ccccgatggc ttcacagacc ccaattgcag ctctgtggtg 120

gaagtggcca gcgacgagga agaacctaca agcgctggcc cctgcacacc caatccatgt 180

cataatggcg gaacctgcga gatcagcgag gcctacagag gcgatacctt catcggctac 240

gtgtgcaagt gccccagagg cttcaatggc atccactgcc agcacaacat caacgagtgc 300

gaggtggaac catgcaagaa cggcggcatc tgtaccgacc tggtggccaa ttactcttgc 360

gagtgccctg gcgagttcat gggcagaaac tgccagtaca aggacgccca caagagcgag 420

gtggcccaca gattcaagga cctgggcgaa gagaacttca aggccctggt gctgatcgcc 480

ttcgctcagt atctccagca gagccctttc gaggaccacg tgaagctggt caacgaagtg 540

accgagttcg ccaagacctg tgtggccgat gagagcgccg agaactgtga caagagcctg 600

cacacactgt tcggcgacaa gctgtgtacc gtggccacac tgagagaaac ctacggcgag 660

atggccgact gctgtgccaa gcaagagccc gagagaaacg agtgcttcct ccagcacaag 720

gatgacaacc ccaacctgcc tagactcgtg cggcctgaag tggatgtgat gtgcaccgcc 780

tttcacgaca acgaggaaac cttcctgaag aagtacctgt acgagatcgc cagacggcac 840

ccctactttt atgcccctga gctgctgttc ttcgccaagc ggtataaggc cgccttcacc 900

gaatgttgcc aggccgctga taaggctgcc tgtctgctgc ctaagctgga cgagctgaga 960

gatgagggca aagccagctc tgccaagcag agactgaaat gcgccagcct ccagaagttc 1020

ggcgagagag cttttaaggc ctgggccgtt gccagactga gccagagatt tcctaaggcc 1080

gagtttgccg aggtgtccaa gctcgtgacc gatctgacaa aggtgcacac cgagtgctgt 1140

cacggcgatc tgctggaatg tgccgacgat agagccgacc tggccaagta tatctgcgag 1200

aaccaggaca gcatcagcag caagctgaaa gagtgctgcg agaagcccct gctggaaaag 1260

tctcactgta tcgccgaagt ggaaaacgac gagatgcccg ccgatctgcc ttctctggct 1320

gccgatttcg tggaaagcaa ggatgtgtgc aagaactacg ccgaggccaa agatgtgttt 1380

ctgggcatgt ttctgtatga gtacgcccgc agacaccccg actattctgt ggttctgctg 1440

ctgcggctgg ccaagacata cgagacaacc ctggaaaaat gctgcgccgc tgccgatcct 1500

cacgagtgtt atgccaaggt gttcgacgag ttcaagccac tggtggaaga accccagaac 1560

ctgatcaagc agaactgcga gctgttcgag cagctgggcg agtacaagtt ccagaatgcc 1620

ctgctcgtgc ggtacaccaa gaaagtgcct caggtgtcca cacctacact ggttgaggtg 1680

tcccggaatc tgggcaaagt gggcagcaag tgttgcaagc accctgaggc caagagaatg 1740

ccttgcgccg aggattacct gagcgtggtg ctgaatcagc tgtgcgtgct gcacgagaaa 1800

acccctgtgt ccgacagagt gaccaagtgc tgtaccgaga gcctcgtgaa cagaaggcct 1860

tgctttagcg ccctggaagt ggacgagaca tacgtgccca aagagttcaa cgccgagaca 1920

ttcaccttcc acgccgatat ctgcaccctg tccgagaaag agcggcagat caagaagcag 1980

acagccctgg tcgagctggt taagcacaag cccaaggcca ccaaagaaca gctgaaggcc 2040

gtgatggacg acttcgccgc ctttgtcgag aagtgctgca aggccgacga caaagagaca 2100

tgcttcgccg aagagggcaa gaaactggtg gcctgtgtgg aacccctcgg catggaaaac 2160

ggcaatatcg ccaatagcca gattgccgcc agcagcgtca gagtgacatt tctgggactg 2220

caacactggg tgcccgagct ggctagactg aatagagccg gcatggtcaa cgcctggaca 2280

cccagcagca acgacgataa cccctggatt caagtgaacc tgctgcggcg tatgtgggtc 2340

acaggtgttg ttacacaggg cgcaagcaga ctggcctctc acgagtacct gaaggccttt 2400

aaggtggcct acagcctgaa cggccacgag ttcgacttca tccacgacgt gaacaagaag 2460

cacaaagagt ttgtcggcaa ctggaacaag aacgccgtgc acgtgaacct gttcgagaca 2520

cctgtggaag cccagtacgt gcggctgtac cctacaagct gtcacaccgc ctgcactctg 2580

agattcgaac tgctgggatg cgagctgaac ggctgtgcta atcctctggg cctgaagaac 2640

aacagcatcc ccgataagca gatcaccgcc agctccagct ataagacatg gggcctgcac 2700

ctgttcagct ggaacccttc ttacgccaga ctggacaagc agggcaactt caatgcttgg 2760

gtggccggca gctacggcaa tgatcagtgg ctgcaagtgg atctgggcag cagcaaagaa 2820

gtgacaggca tcatcaccca aggggccaga aatttcggca gcgtgcagtt cgtggccagc 2880

tacaaagtgg cctactccaa cgacagcgcc aactggaccg agtatcagga ccctagaacc 2940

ggcagctcca agatcttccc cggcaattgg gacaaccaca gccacaagaa gaatctcttc 3000

gagactccca tcctggccag atatgtgcgg attctgcctg tggcctggca caacagaatc 3060

gccctgagac tggaactgct cggctgt 3087

<210> 82

<211> 942

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 82

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu

1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys

20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala

35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn

50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser

65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala

85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu

100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu

115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg

130 135 140

Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg

145 150 155 160

Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn

165 170 175

Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His

180 185 190

Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys

195 200 205

Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu

210 215 220

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

225 230 235 240

Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala

245 250 255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala

260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His

275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala

290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys

305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile
325 330 335
Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala
340 345 350
Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala
355 360 365
Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His
370 375 380
Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu
385 390 395 400
Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr
405 410 415
Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn
420 425 430
Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys
435 440 445
Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val
450 455 460
Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly
465 470 475 480
Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu
485 490 495
Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys
500 505 510
Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val
515 520 525
Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val
530 535 540
Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys
545 550 555 560
Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val

565 570 575

Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala

580 585 590

Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp

595 600 605

Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Cys

610 615 620

Ser Gly Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn Gln Gln Ile

625 630 635 640

Thr Ala Ser Ser Thr His Arg Ala Leu Phe Gly Leu Gln Lys Trp Tyr

645 650 655

Pro Tyr Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp Thr

660 665 670

Ala Ala Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln Arg

675 680 685

Lys Met Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile Gly

690 695 700

Ser Pro Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp Gly

705 710 715 720

Lys Thr Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Asp Met Val

725 730 735

Phe Arg Gly Asn Ile Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe Thr

740 745 750

Pro Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Val Cys Arg

755 760 765

Arg His Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser Gly

770 775 780

Cys Ser Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr Gln

785 790 795 800

Ile Thr Ala Ser Ser Ile Phe Arg Thr Leu Asn Met Asp Met Phe Thr

805 810 815

Trp Glu Pro Arg Lys Ala Arg Leu Asp Lys Gln Gly Lys Val Asn Ala
820 825 830
Trp Thr Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp Leu
835 840 845
Leu Val Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys Asp
850 855 860
Phe Gly His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser Asn
865 870 875 880
Asp Gly Glu His Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys Asp
885 890 895
Lys Val Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn Val
900 905 910
Ile Asp Pro Pro Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp Ser
915 920 925
Trp Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys
930 935 940

<210> 83
<211> 2826
<212> ADN
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"
<400> 83

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60
gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120
aaccactgcg agacaaagga tgcccacaag agcgaggtgg cccacagatt caaggatctg 180
ggcgaagaga acttcaaggc cctggtgctg atcgccttcg ctcagtatct ccagcagagc 240
cctttcgagg accacgtgaa gctggtcaac gaagtgaccg agttcgccaa gacctgtgtg 300
gccgatgaga gcgccgagaa ctgtgataag agcctgcaca ccctgttcgg cgacaagctg 360
tgtacagtgg ccacactgag agaaacctac ggcgagatgg ccgactgctg tgccaagcaa 420

gagcccgaga gaaacgagtg cttcctccag cacaaggacg acaaccccaa cctgcctaga 480

ctcgtgcgac ccgaagtgga tgtgatgtgc accgcctttc acgacaacga ggaaaccttc 540

ctgaagaagt acctgtacga gatcgccaga cggcacccct acttttatgc ccctgagctg 600

ctgttcttcg ccaagcggta taaggccgcc ttcaccgaat gttgccaggc cgctgataag 660

gctgcctgtc tgctgcctaa gctggacgag ctgagagatg agggcaaagc cagctctgcc 720

aagcagagac tgaaatgcgc cagcctccag aagttcggcg agagagcttt taaggcctgg 780

gccgttgcca gactgagcca gagatttcct aaggccgagt ttgccgaggt gtccaagctc 840

gtgaccgatc tgacaaaggt gcacaccgag tgctgtcacg gcgatctgct ggaatgtgcc 900

gacgatagag ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 960

ctgaaagagt gctgcgagaa gcccctgctg gaaaagtctc actgtatcgc cgaggtggaa 1020

aacgacgaga tgcctgccga tctgcctagc ctggctgccg atttcgtgga aagcaaggac 1080

gtgtgcaaga actacgccga ggccaaggat gtgtttctgg gcatgtttct gtatgagtac 1140

gcccgcagac accccgacta ttctgtggtt ctgctgctgc ggctggccaa aacctacgag 1200

acaaccctgg aaaaatgctg cgccgctgcc gatcctcacg agtgttatgc caaggtgttc 1260

gacgagttca agcctctggt ggaagaaccc cagaacctga tcaagcagaa ctgcgagctg 1320

ttcgagcagc tgggcgagta caagttccag aatgccctgc tcgtgcggta caccaagaaa 1380

gtgcctcagg tgtccacacc tacactggtt gaggtgtccc ggaatctggg caaagtgggc 1440

agcaagtgtt gcaagcaccc tgaggccaag agaatgcctt gcgccgagga ttacctgagc 1500

gtggtgctga atcagctgtg cgtgctgcac gagaaaaccc ctgtgtccga cagagtgacc 1560

aagtgctgta ccgagagcct cgtgaacaga aggccttgct ttagcgccct ggaagtggac 1620

gagacatacg tgcccaaaga gttcaacgcc gagacattca ccttccacgc cgacatctgc 1680

accctgtccg agaaagagcg gcagatcaag aagcagacag ccctggtcga gctggttaag 1740

cacaagccca aggccaccaa agaacagctg aaggccgtga tggacgactt cgccgccttt 1800

gtcgagaagt gctgcaaggc cgacgacaaa gagacatgct tcgccgaaga gggcaagaaa 1860

ctggtggcct gttctggccc tctgggcatc gaaggcggca tcatcagcaa tcagcagatc 1920

accgccagca gcacccacag agcactgttt ggcctgcaaa agtggtatcc ctactacgcc 1980

cggctgaaca agaagggcct gattaacgcc tggacagccg ccgagaatga cagatggccc 2040

tggattcaga tcaacctcca gcggaagatg agagtgaccg gcgttatcac acagggcgca 2100

aagagaatcg gctcccctga gtacatcaag agctacaaga tcgcctacag caacgacggc 2160

aagacctggg ccatgtacaa agtgaagggc accaacgagg acatggtgtt ccggggcaac 2220

atcgacaaca acacccctta cgccaacagc ttcacccctc ctatcaaggc ccagtacgtg 2280

cggctgtacc ctcaagtgtg cagaaggcac tgtaccctga gaatggaact gctgggctgc 2340

gaactgtctg gctgttctga gccactggga atgaagtccg gccacatcca ggactaccag 2400

attaccgcct ccagcatctt cagaaccctg aacatggata tgttcacctg ggagccccgg 2460

aaggccagac tggataagca gggaaaagtg aatgcctgga ccagcggcca caacgaccag 2520

tctcaatggc tgcaagtgga cctgctggtg cctaccaaag tgaccggaat catcacccaa 2580

ggcgctaagg atttcggcca cgtgcagttc gtgggctcct acaagctggc ctactccaat 2640

gatggcgagc actggaccgt gtaccaggac gagaagcagc ggaaggataa ggtgttccag 2700

ggaaacttcg ataacgatac ccaccggaag aacgtgatcg accctccaat ctacgccaga 2760

cacatcagaa tcctgccttg gtcttggtac ggcagaatca ccctgagatc cgagctgctg 2820

ggatgc 2826

<210> 84

<211> 968

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 84

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45

Pro Thr Gly Ser Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys
50 55 60

Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala
65 70 75 80

Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn
85 90 95

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

100 105 110

Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr

115 120 125

Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala

130 135 140

Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp

145 150 155 160

Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys

165 170 175

Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr

180 185 190

Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe

195 200 205

Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala

210 215 220

Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu

225 230 235 240

Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln

245 250 255

Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser

260 265 270

Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr

275 280 285

Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu

290 295 300

Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln

305 310 315 320

Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu

325 330 335

Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala

340 345 350

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

595 600 605

Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly

610 615 620

Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser

625 630 635 640

Gly Gly Ser Gly Gly Ser Gly Gly Ser Cys Ser Gly Pro Leu Gly Ile

645 650 655

Glu Gly Gly Ile Ile Ser Asn Gln Gln Ile Thr Ala Ser Ser Thr His

660 665 670

Arg Ala Leu Phe Gly Leu Gln Lys Trp Tyr Pro Tyr Tyr Ala Arg Leu

675 680 685

Asn Lys Lys Gly Leu Ile Asn Ala Trp Thr Ala Ala Glu Asn Asp Arg

690 695 700

Trp Pro Trp Ile Gln Ile Asn Leu Gln Arg Lys Met Arg Val Thr Gly

705 710 715 720

Val Ile Thr Gln Gly Ala Lys Arg Ile Gly Ser Pro Glu Tyr Ile Lys

725 730 735

Ser Tyr Lys Ile Ala Tyr Ser Asn Asp Gly Lys Thr Trp Ala Met Tyr

740 745 750

Lys Val Lys Gly Thr Asn Glu Asp Met Val Phe Arg Gly Asn Ile Asp

755 760 765

Asn Asn Thr Pro Tyr Ala Asn Ser Phe Thr Pro Pro Ile Lys Ala Gln

770 775 780

Tyr Val Arg Leu Tyr Pro Gln Val Cys Arg Arg His Cys Thr Leu Arg

785 790 795 800

Met Glu Leu Leu Gly Cys Glu Leu Ser Gly Cys Ser Glu Pro Leu Gly

805 810 815

Met Lys Ser Gly His Ile Gln Asp Tyr Gln Ile Thr Ala Ser Ser Ile

820 825 830

Phe Arg Thr Leu Asn Met Asp Met Phe Thr Trp Glu Pro Arg Lys Ala

835 840 845

Arg Leu Asp Lys Gln Gly Lys Val Asn Ala Trp Thr Ser Gly His Asn
850 855 860

Asp Gln Ser Gln Trp Leu Gln Val Asp Leu Leu Val Pro Thr Lys Val
865 870 875 880

Thr Gly Ile Ile Thr Gln Gly Ala Lys Asp Phe Gly His Val Gln Phe
885 890 895

Val Gly Ser Tyr Lys Leu Ala Tyr Ser Asn Asp Gly Glu His Trp Thr
900 905 910

Val Tyr Gln Asp Glu Lys Gln Arg Lys Asp Lys Val Phe Gln Gly Asn
915 920 925

Phe Asp Asn Asp Thr His Arg Lys Asn Val Ile Asp Pro Pro Ile Tyr
930 935 940

Ala Arg His Ile Arg Ile Leu Pro Trp Ser Trp Tyr Gly Arg Ile Thr
945 950 955 960

Leu Arg Ser Glu Leu Leu Gly Cys
965

<210> 85

<211> 963

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 85

Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly Thr Cys
1 5 10 15

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

Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Gly Ser Asp
35 40 45

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

50 55 60

Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln

65 70 75 80

Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe

85 90 95

Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser

100 105 110

Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg

115 120 125

Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu

130 135 140

Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro

145 150 155 160

Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp

165 170 175

Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg

180 185 190

His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr

195 200 205

Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys

210 215 220

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

225 230 235 240

Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg

245 250 255

Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys

260 265 270

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

275 280 285

His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg

290 295 300

Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser
305 310 315 320
Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys
325 330 335
Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu
340 345 350
Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu
355 360 365
Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg
370 375 380
His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr
385 390 395 400
Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys
405 410 415
Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln
420 425 430
Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr
435 440 445
Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln
450 455 460
Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val
465 470 475 480
Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala
485 490 495
Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu
500 505 510
Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu
515 520 525
Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr
530 535 540
Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile

545 550 555 560

Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu

565 570 575

Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys

580 585 590

Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala

595 600 605

Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala

610 615 620

Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly Gly

625 630 635 640

Ser Gly Gly Ser Cys Ser Gly Pro Leu Gly Ile Glu Gly Gly Ile Ile

645 650 655

Ser Asn Gln Gln Ile Thr Ala Ser Ser Thr His Arg Ala Leu Phe Gly

660 665 670

Leu Gln Lys Trp Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys Lys Gly Leu

675 680 685

Ile Asn Ala Trp Thr Ala Ala Glu Asn Asp Arg Trp Pro Trp Ile Gln

690 695 700

Ile Asn Leu Gln Arg Lys Met Arg Val Thr Gly Val Ile Thr Gln Gly

705 710 715 720

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

725 730 735

Tyr Ser Asn Asp Gly Lys Thr Trp Ala Met Tyr Lys Val Lys Gly Thr

740 745 750

Asn Glu Asp Met Val Phe Arg Gly Asn Ile Asp Asn Asn Thr Pro Tyr

755 760 765

Ala Asn Ser Phe Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr

770 775 780

Pro Gln Val Cys Arg Arg His Cys Thr Leu Arg Met Glu Leu Leu Gly

785 790 795 800

Cys Glu Leu Ser Gly Cys Ser Glu Pro Leu Gly Met Lys Ser Gly His
805 810 815
Ile Gln Asp Tyr Gln Ile Thr Ala Ser Ser Ile Phe Arg Thr Leu Asn
820 825 830
Met Asp Met Phe Thr Trp Glu Pro Arg Lys Ala Arg Leu Asp Lys Gln
835 840 845
Gly Lys Val Asn Ala Trp Thr Ser Gly His Asn Asp Gln Ser Gln Trp
850 855 860
Leu Gln Val Asp Leu Leu Val Pro Thr Lys Val Thr Gly Ile Ile Thr
865 870 875 880
Gln Gly Ala Lys Asp Phe Gly His Val Gln Phe Val Gly Ser Tyr Lys
885 890 895
Leu Ala Tyr Ser Asn Asp Gly Glu His Trp Thr Val Tyr Gln Asp Glu
900 905 910
Lys Gln Arg Lys Asp Lys Val Phe Gln Gly Asn Phe Asp Asn Asp Thr
915 920 925
His Arg Lys Asn Val Ile Asp Pro Pro Ile Tyr Ala Arg His Ile Arg
930 935 940
Ile Leu Pro Trp Ser Trp Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu
945 950 955 960
Leu Gly Cys

<210> 86
<211> 957
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 86
Asn Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys

1 5 10 15

Thr Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met

20 25 30

Gly Arg Asn Cys Gln Tyr Lys Gly Ser Asp Ala His Lys Ser Glu Val

35 40 45

Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val

50 55 60

Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His

65 70 75 80

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

85 90 95

Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly

100 105 110

Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met

115 120 125

Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu

130 135 140

Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu

145 150 155 160

Val Asp Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu

165 170 175

Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala

180 185 190

Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu

195 200 205

Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp

210 215 220

Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys

225 230 235 240

Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala

245 250 255

Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val
260 265 270
Ser Lys Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His
275 280 285
Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr
290 295 300
Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys
305 310 315 320
Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn
325 330 335
Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu
340 345 350
Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu
355 360 365
Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val
370 375 380
Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys
385 390 395 400
Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp
405 410 415
Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn
420 425 430
Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu
435 440 445
Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu
450 455 460
Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys
465 470 475 480
His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val
485 490 495
Val Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp

500 505 510

Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys

515 520 525

Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn

530 535 540

Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys

545 550 555 560

Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His

565 570 575

Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe

580 585 590

Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys

595 600 605

Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu

610 615 620

Gly Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Cys Ser

625 630 635 640

Gly Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn Gln Gln Ile Thr

645 650 655

Ala Ser Ser Thr His Arg Ala Leu Phe Gly Leu Gln Lys Trp Tyr Pro

660 665 670

Tyr Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp Thr Ala

675 680 685

Ala Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln Arg Lys

690 695 700

Met Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile Gly Ser

705 710 715 720

Pro Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp Gly Lys

725 730 735

Thr Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Asp Met Val Phe

740 745 750

Arg Gly Asn Ile Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe Thr Pro
755 760 765

Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Val Cys Arg Arg
770 775 780

His Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser Gly Cys
785 790 795 800

Ser Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr Gln Ile
805 810 815

Thr Ala Ser Ser Ile Phe Arg Thr Leu Asn Met Asp Met Phe Thr Trp
820 825 830

Glu Pro Arg Lys Ala Arg Leu Asp Lys Gln Gly Lys Val Asn Ala Trp
835 840 845

Thr Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp Leu Leu
850 855 860

Val Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys Asp Phe
865 870 875 880

Gly His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser Asn Asp
885 890 895

Gly Glu His Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys Asp Lys
900 905 910

Val Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn Val Ile
915 920 925

Asp Pro Pro Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp Ser Trp
930 935 940

Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys
945 950 955

<210> 87

<211> 1013

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 87

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45

Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly
50 55 60

Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr
65 70 75 80

Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Gly
85 90 95

Ser Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly
100 105 110

Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu
115 120 125

Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr
130 135 140

Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp
145 150 155 160

Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr
165 170 175

Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu
180 185 190

Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn
195 200 205

Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe
210 215 220

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

465 470 475 480

Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly

485 490 495

Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val

500 505 510

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

515 520 525

Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro

530 535 540

Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu

545 550 555 560

His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu

565 570 575

Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu

580 585 590

Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala

595 600 605

Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr

610 615 620

Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln

625 630 635 640

Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys

645 650 655

Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu

660 665 670

Val Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser

675 680 685

Gly Gly Ser Gly Gly Ser Cys Ser Gly Pro Leu Gly Ile Glu Gly Gly

690 695 700

Ile Ile Ser Asn Gln Gln Ile Thr Ala Ser Ser Thr His Arg Ala Leu

705 710 715 720

Phe Gly Leu Gln Lys Trp Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys Lys
725 730 735
Gly Leu Ile Asn Ala Trp Thr Ala Ala Glu Asn Asp Arg Trp Pro Trp
740 745 750
Ile Gln Ile Asn Leu Gln Arg Lys Met Arg Val Thr Gly Val Ile Thr
755 760 765
Gln Gly Ala Lys Arg Ile Gly Ser Pro Glu Tyr Ile Lys Ser Tyr Lys
770 775 780
Ile Ala Tyr Ser Asn Asp Gly Lys Thr Trp Ala Met Tyr Lys Val Lys
785 790 795 800
Gly Thr Asn Glu Asp Met Val Phe Arg Gly Asn Ile Asp Asn Asn Thr
805 810 815
Pro Tyr Ala Asn Ser Phe Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg
820 825 830
Leu Tyr Pro Gln Val Cys Arg Arg His Cys Thr Leu Arg Met Glu Leu
835 840 845
Leu Gly Cys Glu Leu Ser Gly Cys Ser Glu Pro Leu Gly Met Lys Ser
850 855 860
Gly His Ile Gln Asp Tyr Gln Ile Thr Ala Ser Ser Ile Phe Arg Thr
865 870 875 880
Leu Asn Met Asp Met Phe Thr Trp Glu Pro Arg Lys Ala Arg Leu Asp
885 890 895
Lys Gln Gly Lys Val Asn Ala Trp Thr Ser Gly His Asn Asp Gln Ser
900 905 910
Gln Trp Leu Gln Val Asp Leu Leu Val Pro Thr Lys Val Thr Gly Ile
915 920 925
Ile Thr Gln Gly Ala Lys Asp Phe Gly His Val Gln Phe Val Gly Ser
930 935 940
Tyr Lys Leu Ala Tyr Ser Asn Asp Gly Glu His Trp Thr Val Tyr Gln
945 950 955 960
Asp Glu Lys Gln Arg Lys Asp Lys Val Phe Gln Gly Asn Phe Asp Asn

965 970 975

Asp Thr His Arg Lys Asn Val Ile Asp Pro Pro Ile Tyr Ala Arg His

980 985 990

Ile Arg Ile Leu Pro Trp Ser Trp Tyr Gly Arg Ile Thr Leu Arg Ser

995 1000 1005

Glu Leu Leu Gly Cys

1010

<210> 88

<211> 1002

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 88

Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly Thr Cys

1 5 10 15

Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr Val Cys

20 25 30

Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn Ile Asn

35 40 45

Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr Asp Leu

50 55 60

Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly Arg Asn

65 70 75 80

Cys Gln Tyr Lys Gly Ser Asp Ala His Lys Ser Glu Val Ala His Arg

85 90 95

Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala

100 105 110

Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu

115 120 125

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

370 375 380

Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp

385 390 395 400

Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe

405 410 415

Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu

420 425 430

Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala

435 440 445

Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys

450 455 460

Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu

465 470 475 480

Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg

485 490 495

Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val

500 505 510

Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu

515 520 525

Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn

530 535 540

Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr

545 550 555 560

Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala

565 570 575

Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr

580 585 590

Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln

595 600 605

Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys

610 615 620

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

865 870 875 880

Lys Ala Arg Leu Asp Lys Gln Gly Lys Val Asn Ala Trp Thr Ser Gly

885 890 895

His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp Leu Leu Val Pro Thr

900 905 910

Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys Asp Phe Gly His Val

915 920 925

Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser Asn Asp Gly Glu His

930 935 940

Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys Asp Lys Val Phe Gln

945 950 955 960

Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn Val Ile Asp Pro Pro

965 970 975

Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp Ser Trp Tyr Gly Arg

980 985 990

Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys

995 1000

<210> 89

<211> 1007

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 89

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro

1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr

20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu

35 40 45

Pro Thr Asn Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly
50 55 60
Ile Cys Thr Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu
65 70 75 80
Phe Met Gly Arg Asn Cys Gln Tyr Lys Gly Ser Asp Ala His Lys Ser
85 90 95
Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala
100 105 110
Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu
115 120 125
Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys
130 135 140
Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu
145 150 155 160
Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly
165 170 175
Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys
180 185 190
Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg
195 200 205
Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr
210 215 220
Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe
225 230 235 240
Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe
245 250 255
Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys
260 265 270
Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg
275 280 285
Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala

290 295 300

Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala

305 310 315 320

Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys

325 330 335

Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala

340 345 350

Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu

355 360 365

Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val

370 375 380

Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe

385 390 395 400

Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val

405 410 415

Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr

420 425 430

Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu

435 440 445

Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val

450 455 460

Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys

465 470 475 480

Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn

485 490 495

Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro

500 505 510

Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys

515 520 525

Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu

530 535 540

Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val
545 550 555 560
Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg
565 570 575
Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu
580 585 590
Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser
595 600 605
Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val
610 615 620
Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp
625 630 635 640
Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu
645 650 655
Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala
660 665 670
Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser
675 680 685
Cys Ser Gly Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn Gln Gln
690 695 700
Ile Thr Ala Ser Ser Thr His Arg Ala Leu Phe Gly Leu Gln Lys Trp
705 710 715 720
Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp
725 730 735
Thr Ala Ala Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln
740 745 750
Arg Lys Met Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile
755 760 765
Gly Ser Pro Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp
770 775 780
Gly Lys Thr Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Asp Met

785 790 795 800

Val Phe Arg Gly Asn Ile Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe

805 810 815

Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Val Cys

820 825 830

Arg Arg His Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser

835 840 845

Gly Cys Ser Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr

850 855 860

Gln Ile Thr Ala Ser Ser Ile Phe Arg Thr Leu Asn Met Asp Met Phe

865 870 875 880

Thr Trp Glu Pro Arg Lys Ala Arg Leu Asp Lys Gln Gly Lys Val Asn

885 890 895

Ala Trp Thr Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp

900 905 910

Leu Leu Val Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys

915 920 925

Asp Phe Gly His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser

930 935 940

Asn Asp Gly Glu His Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys

945 950 955 960

Asp Lys Val Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn

965 970 975

Val Ile Asp Pro Pro Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp

980 985 990

Ser Trp Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys

995 1000 1005

<210> 90

<211> 2904

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 90

gacatctgcg accccaaccc ctgcgagaac ggcggcatct gcctgcccgg cctggccgac 60

ggcagcttca gctgcgagtg ccccgacggc ttcaccgacc ccaactgcag cagcgtggtg 120

gaggtggcca gcgacgagga ggagcccacc ggcagcgacg cccacaagag cgaggtggcc 180

caccggttca aggacctggg cgaggagaac ttcaaggccc tggtgctgat cgccttcgcc 240

cagtacctgc agcagagccc cttcgaggac cacgtgaagc tggtgaacga ggtgaccgag 300

ttcgccaaga cctgcgtggc cgacgagagc gccgagaact gcgacaagag cctgcacacc 360

ctgttcggcg acaagctgtg caccgtggcc accctgcggg agacctacgg cgagatggcc 420

gactgctgcg ccaagcagga gcccgagcgg aacgagtgct tcctgcagca caaggacgac 480

aaccccaacc tgccccggct ggtgcggccc gaggtggacg tgatgtgcac cgccttccac 540

gacaacgagg agaccttcct gaagaagtac ctgtacgaga tcgcccggcg gcacccctac 600

ttctacgccc ccgagctgct gttcttcgcc aagcggtaca aggccgcctt caccgagtgc 660

tgccaggccg ccgacaaggc cgcctgcctg ctgcccaagc tggacgagct gcgggacgag 720

ggcaaggcca gcagcgccaa gcagcggctg aagtgcgcca gcctgcagaa gttcggcgag 780

cgggccttca aggcctgggc cgtggcccgg ctgagccagc ggttccccaa ggccgagttc 840

gccgaggtga gcaagctggt gaccgacctg accaaggtgc acaccgagtg ctgccacggc 900

gacctgctgg agtgcgccga cgaccgggcc gacctggcca agtacatctg cgagaaccag 960

gacagcatca gcagcaagct gaaggagtgc tgcgagaagc ccctgctgga gaagagccac 1020

tgcatcgccg aggtggagaa cgacgagatg cccgccgacc tgcccagcct ggccgccgac 1080

ttcgtggaga gcaaggacgt gtgcaagaac tacgccgagg ccaaggacgt gttcctgggc 1140

atgttcctgt acgagtacgc ccggcggcac cccgactaca gcgtggtgct gctgctgcgg 1200

ctggccaaga cctacgagac caccctggag aagtgctgcg ccgccgccga cccccacgag 1260

tgctacgcca aggtgttcga cgagttcaag cccctggtgg aggagcccca gaacctgatc 1320

aagcagaact gcgagctgtt cgagcagctg ggcgagtaca agttccagaa cgccctgctg 1380

gtgcggtaca ccaagaaggt gccccaggtg agcaccccca ccctggtgga ggtgagccgg 1440

aacctgggca aggtgggcag caagtgctgc aagcaccccg aggccaagcg gatgccctgc 1500

gccgaggact acctgagcgt ggtgctgaac cagctgtgcg tgctgcacga gaagaccccc 1560

gtgagcgacc gggtgaccaa gtgctgcacc gagagcctgg tgaaccggcg gccctgcttc 1620

agcgccctgg aggtggacga gacctacgtg cccaaggagt tcaacgccga gaccttcacc 1680

ttccacgccg acatctgcac cctgagcgag aaggagcggc agatcaagaa gcagaccgcc 1740

ctggtggagc tggtgaagca caagcccaag gccaccaagg agcagctgaa ggccgtgatg 1800

gacgacttcg ccgccttcgt ggagaagtgc tgcaaggccg acgacaagga gacctgcttc 1860

gccgaggagg gcaagaagct ggtggccgcc agccaggccg ccctgggcct gggcggcagc 1920

ggcggcagcg gcggcagcgg cggcagctgc agcggccccc tgggcatcga gggcggcatc 1980

atcagcaacc agcagatcac cgccagcagc acccaccggg ccctgttcgg cctgcagaag 2040

tggtacccct actacgcccg gctgaacaag aagggcctga tcaacgcctg gaccgccgcc 2100

gagaacgacc ggtggccctg gatccagatc aacctgcagc ggaagatgcg ggtgaccggc 2160

gtgatcaccc agggcgccaa gcggatcggc agccccgagt acatcaagag ctacaagatc 2220

gcctacagca acgacggcaa gacctgggcc atgtacaagg tgaagggcac caacgaggac 2280

atggtgttcc ggggcaacat cgacaacaac acccccctacg ccaacagctt caccccccccc 2340

atcaaggccc agtacgtgcg gctgtacccc caggtgtgcc ggcggcactg caccctgcgg 2400

atggagctgc tgggctgcga gctgagcggc tgcagcgagc ccctgggcat gaagagcggc 2460

cacatccagg actaccagat caccgccagc agcatcttcc ggaccctgaa catggacatg 2520

ttcacctggg agccccggaa ggcccggctg gacaagcagg gcaaggtgaa cgcctggacc 2580

agcggccaca acgaccagag ccagtggctg caggtggacc tgctggtgcc caccaaggtg 2640

accggcatca tcacccaggg cgccaaggac ttcggccacg tgcagttcgt gggcagctac 2700

aagctggcct acagcaacga cggcgagcac tggaccgtgt accaggacga gaagcagcgg 2760

aaggacaagg tgttccaggg caacttcgac aacgacaccc accggaagaa cgtgatcgac 2820

ccccccatct acgcccggca catccggatc ctgccctgga gctggtacgg ccggatcacc 2880

ctgcggagcg agctgctggg ctgc 2904

<210> 91

<211> 2889

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 91

agcgccggcc cctgcacccc caacccctgc cacaacggcg gcacctgcga gatcagcgag 60

gcctaccggg gcgacacctt catcggctac gtgtgcaagt gcccccgggg cttcaacggc 120

atccactgcc agcacggcag cgacgcccac aagagcgagg tggcccaccg gttcaaggac 180

ctgggcgagg agaacttcaa ggccctggtg ctgatcgcct tcgcccagta cctgcagcag 240

agcccccttcg aggaccacgt gaagctggtg aacgaggtga ccgagttcgc caagacctgc 300

gtggccgacg agagcgccga gaactgcgac aagagcctgc acaccctgtt cggcgacaag 360

ctgtgcaccg tggccaccct gcgggagacc tacggcgaga tggccgactg ctgcgccaag 420

caggagcccg agcggaacga gtgcttcctg cagcacaagg acgacaaccc caacctgccc 480

cggctggtgc ggcccgaggt ggacgtgatg tgcaccgcct tccacgacaa cgaggagacc 540

ttcctgaaga agtacctgta cgagatcgcc cggcggcacc cctacttcta cgcccccgag 600

ctgctgttct tcgccaagcg gtacaaggcc gccttcaccg agtgctgcca ggccgccgac 660

aaggccgcct gcctgctgcc caagctggac gagctgcggg acgagggcaa ggccagcagc 720

gccaagcagc ggctgaagtg cgccagcctg cagaagttcg gcgagcgggc cttcaaggcc 780

tgggccgtgg cccggctgag ccagcggttc cccaaggccg agttcgccga ggtgagcaag 840

ctggtgaccg acctgaccaa ggtgcacacc gagtgctgcc acggcgacct gctggagtgc 900

gccgacgacc gggccgacct ggccaagtac atctgcgaga accaggacag catcagcagc 960

aagctgaagg agtgctgcga gaagcccctg ctggagaaga gccactgcat cgccgaggtg 1020

gagaacgacg agatgcccgc cgacctgccc agcctggccg ccgacttcgt ggagagcaag 1080

gacgtgtgca agaactacgc cgaggccaag gacgtgttcc tgggcatgtt cctgtacgag 1140

tacgcccggc ggcaccccga ctacagcgtg gtgctgctgc tgcggctggc caagacctac 1200

gagaccaccc tggagaagtg ctgcgccgcc gccgaccccc acgagtgcta cgccaaggtg 1260

ttcgacgagt tcaagccccct ggtggaggag ccccagaacc tgatcaagca gaactgcgag 1320

ctgttcgagc agctgggcga gtacaagttc cagaacgccc tgctggtgcg gtacaccaag 1380

aaggtgcccc aggtgagcac ccccaccctg gtggaggtga gccggaacct gggcaaggtg 1440

ggcagcaagt gctgcaagca ccccgaggcc aagcggatgc cctgcgccga ggactacctg 1500

agcgtggtgc tgaaccagct gtgcgtgctg cacgagaaga cccccgtgag cgaccgggtg 1560

accaagtgct gcaccgagag cctggtgaac cggcggccct gcttcagcgc cctggaggtg 1620

gacgagacct acgtgcccaa ggagttcaac gccgagacct tcaccttcca cgccgacatc 1680

tgcaccctga gcgagaagga gcggcagatc aagaagcaga ccgccctggt ggagctggtg 1740

aagcacaagc ccaaggccac caaggagcag ctgaaggccg tgatggacga cttcgccgcc 1800

ttcgtggaga agtgctgcaa ggccgacgac aaggagacct gcttcgccga ggagggcaag 1860

aagctggtgg ccgccagcca ggccgccctg ggcctgggcg gcagcggcgg cagcggcggc 1920

agcggcggca gctgcagcgg ccccctgggc atcgagggcg gcatcatcag caaccagcag 1980

atcaccgcca gcagcaccca ccgggccctg ttcggcctgc agaagtggta cccctactac 2040

gcccggctga acaagaaggg cctgatcaac gcctggaccg ccgccgagaa cgaccggtgg 2100

ccctggatcc agatcaacct gcagcggaag atgcgggtga ccggcgtgat cacccagggc 2160

gccaagcgga tcggcagccc cgagtacatc aagagctaca agatcgccta cagcaacgac 2220

ggcaagacct gggccatgta caaggtgaag ggcaccaacg aggacatggt gttccggggc 2280

aacatcgaca acaacacccc ctacgccaac agcttcaccc cccccatcaa ggcccagtac 2340

gtgcggctgt acccccaggt gtgccggcgg cactgcaccc tgcggatgga gctgctgggc 2400

tgcgagctga gcggctgcag cgagcccctg ggcatgaaga gcggccacat ccaggactac 2460

cagatcaccg ccagcagcat cttccggacc ctgaacatgg acatgttcac ctgggagccc 2520

cggaaggccc ggctggacaa gcagggcaag gtgaacgcct ggaccagcgg ccacaacgac 2580

cagagccagt ggctgcaggt ggacctgctg gtgcccacca aggtgaccgg catcatcacc 2640

cagggcgcca aggacttcgg ccacgtgcag ttcgtgggca gctacaagct ggcctacagc 2700

aacgacggcg agcactggac cgtgtaccag gacgagaagc agcggaagga caaggtgttc 2760

cagggcaact tcgacaacga cacccaccgg aagaacgtga tcgacccccc catctacgcc 2820

cggcacatcc ggatcctgcc ctggagctgg tacggccgga tcaccctgcg gagcgagctg 2880

ctgggctgc 2889

<210> 92

<211> 2871

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 92

aacatcaacg agtgcgaggt ggagccctgc aagaacggcg gcatctgcac cgacctggtg 60

gccaactaca gctgcgagtg ccccggcgag ttcatgggcc ggaactgcca gtacaagggc 120

agcgacgccc acaagagcga ggtggcccac cggttcaagg acctgggcga ggagaacttc 180

aaggccctgg tgctgatcgc cttcgcccag tacctgcagc agagcccctt cgaggaccac 240

gtgaagctgg tgaacgaggt gaccgagttc gccaagacct gcgtggccga cgagagcgcc 300

gagaactgcg acaagagcct gcacaccctg ttcggcgaca agctgtgcac cgtggccacc 360

ctgcgggaga cctacggcga gatggccgac tgctgcgcca agcaggagcc cgagcggaac 420

gagtgcttcc tgcagcacaa ggacgacaac cccaacctgc cccggctggt gcggcccgag 480

gtggacgtga tgtgcaccgc cttccacgac aacgaggaga ccttcctgaa gaagtacctg 540

tacgagatcg cccggcggca cccctacttc tacgcccccg agctgctgtt cttcgccaag 600

cggtacaagg ccgccttcac cgagtgctgc caggccgccg acaaggccgc ctgcctgctg 660

cccaagctgg acgagctgcg ggacgagggc aaggccagca gcgccaagca gcggctgaag 720

tgcgccagcc tgcagaagtt cggcgagcgg gccttcaagg cctgggccgt ggcccggctg 780

agccagcggt tccccaaggc cgagttcgcc gaggtgagca agctggtgac cgacctgacc 840

aaggtgcaca ccgagtgctg ccacggcgac ctgctggagt gcgccgacga ccgggccgac 900

ctggccaagt acatctgcga gaaccaggac agcatcagca gcaagctgaa ggagtgctgc 960

gagaagcccc tgctggagaa gagccactgc atcgccgagg tggagaacga cgagatgccc 1020

gccgacctgc ccagcctggc cgccgacttc gtggagagca aggacgtgtg caagaactac 1080

gccgaggcca aggacgtgtt cctgggcatg ttcctgtacg agtacgcccg gcggcacccc 1140

gactacagcg tggtgctgct gctgcggctg gccaagacct acgagaccac cctggagaag 1200

tgctgcgccg ccgccgaccc ccacgagtgc tacgccaagg tgttcgacga gttcaagccc 1260

ctggtggagg agccccagaa cctgatcaag cagaactgcg agctgttcga gcagctgggc 1320

gagtacaagt tccagaacgc cctgctggtg cggtacacca agaaggtgcc ccaggtgagc 1380

acccccaccc tggtggaggt gagccggaac ctgggcaagg tgggcagcaa gtgctgcaag 1440

caccccgagg ccaagcggat gccctgcgcc gaggactacc tgagcgtggt gctgaaccag 1500

ctgtgcgtgc tgcacgagaa gacccccgtg agcgaccggg tgaccaagtg ctgcaccgag 1560

agcctggtga accggcggcc ctgcttcagc gccctggagg tggacgagac ctacgtgccc 1620

aaggagttca acgccgagac cttcaccttc cacgccgaca tctgcaccct gagcgagaag 1680

gagcggcaga tcaagaagca gaccgccctg gtggagctgg tgaagcacaa gcccaaggcc 1740

accaaggagc agctgaaggc cgtgatggac gacttcgccg ccttcgtgga gaagtgctgc 1800

aaggccgacg acaaggagac ctgcttcgcc gaggagggca agaagctggt ggccgccagc 1860

caggccgccc tgggcctggg cggcagcggc ggcagcggcg gcagcggcgg cagctgcagc 1920

ggcccccctgg gcatcgaggg cggcatcatc agcaaccagc agatcaccgc cagcagcacc 1980

caccgggccc tgttcggcct gcagaagtgg taccctact acgcccggct gaacaagaag 2040

ggcctgatca acgcctggac cgccgccgag aacgaccggt ggccctggat ccagatcaac 2100

ctgcagcgga agatgcgggt gaccggcgtg atcacccagg gcgccaagcg gatcggcagc 2160

cccgagtaca tcaagagcta caagatcgcc tacagcaacg acggcaagac ctgggccatg 2220

tacaaggtga agggcaccaa cgaggacatg gtgttccggg gcaacatcga caacaacacc 2280

ccctacgcca acagcttcac ccccccccatc aaggcccagt acgtgcggct gtaccccccag 2340

gtgtgccggc ggcactgcac cctgcggatg gagctgctgg gctgcgagct gagcggctgc 2400

agcgagcccc tgggcatgaa gagcggccac atccaggact accagatcac cgccagcagc 2460

atcttccgga ccctgaacat ggacatgttc acctgggagc cccggaaggc ccggctggac 2520

aagcagggca aggtgaacgc ctggaccagc ggccacaacg accagagcca gtggctgcag 2580

gtggacctgc tggtgcccac caaggtgacc ggcatcatca cccagggcgc caaggacttc 2640

ggccacgtgc agttcgtggg cagctacaag ctggcctaca gcaacgacgg cgagcactgg 2700

accgtgtacc aggacgagaa gcagcggaag gacaaggtgt tccagggcaa cttcgacaac 2760

gacacccacc ggaagaacgt gatcgacccc cccatctacg cccggcacat ccggatcctg 2820

ccctggagct ggtacggccg gatcaccctg cggagcgagc tgctgggctg c 2871

<210> 93

<211> 3039

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 93

gacatctgcg accccaaccc ctgcgagaac ggcggcatct gcctgcccgg cctggccgac 60

ggcagcttca gctgcgagtg ccccgacggc ttcaccgacc ccaactgcag cagcgtggtg 120

gaggtggcca gcgacgagga ggagcccacc agcgccggcc cctgcacccc caacccctgc 180

cacaacggcg gcacctgcga gatcagcgag gcctaccggg gcgacacctt catcggctac 240

gtgtgcaagt gccccgggg cttcaacggc atccactgcc agcacggcag cgacgcccac 300

aagagcgagg tggcccaccg gttcaaggac ctgggcgagg agaacttcaa ggccctggtg 360

ctgatcgcct tcgcccagta cctgcagcag agcccttcg aggaccacgt gaagctggtg 420

aacgaggtga ccgagttcgc caagacctgc gtggccgacg agagcgccga gaactgcgac 480

aagagcctgc acaccctgtt cggcgacaag ctgtgcaccg tggccaccct gcgggagacc 540

tacggcgaga tggccgactg ctgcgccaag caggagcccg agcggaacga gtgcttcctg 600

cagcacaagg acgacaaccc caacctgccc cggctggtgc ggcccgaggt ggacgtgatg 660

tgcaccgcct tccacgacaa cgaggagacc ttcctgaaga agtacctgta cgagatcgcc 720

cggcggcacc cctacttcta cgcccccgag ctgctgttct tcgccaagcg gtacaaggcc 780

gccttcaccg agtgctgcca ggccgccgac aaggccgcct gcctgctgcc caagctggac 840

gagctgcggg acgagggcaa ggccagcagc gccaagcagc ggctgaagtg cgccagcctg 900

cagaagttcg gcgagcgggc cttcaaggcc tgggccgtgg cccggctgag ccagcggttc 960

cccaaggccg agttcgccga ggtgagcaag ctggtgaccg acctgaccaa ggtgcacacc 1020

gagtgctgcc acggcgacct gctggagtgc gccgacgacc gggccgacct ggccaagtac 1080

atctgcgaga accaggacag catcagcagc aagctgaagg agtgctgcga gaagcccctg 1140

ctggagaaga gccactgcat cgccgaggtg gagaacgacg agatgcccgc cgacctgccc 1200

agcctggccg ccgacttcgt ggagagcaag gacgtgtgca agaactacgc cgaggccaag 1260

gacgtgttcc tgggcatgtt cctgtacgag tacgcccggc ggcaccccga ctacagcgtg 1320

gtgctgctgc tgcggctggc caagacctac gagaccaccc tggagaagtg ctgcgccgcc 1380

gccgaccccc acgagtgcta cgccaaggtg ttcgacgagt tcaagccccc ggtggaggag 1440

ccccagaacc tgatcaagca gaactgcgag ctgttcgagc agctgggcga gtacaagttc 1500

cagaacgccc tgctggtgcg gtacaccaag aaggtgcccc aggtgagcac ccccaccctg 1560

gtggaggtga gccggaacct gggcaaggtg ggcagcaagt gctgcaagca ccccgaggcc 1620

aagcggatgc cctgcgccga ggactacctg agcgtggtgc tgaaccagct gtgcgtgctg 1680

cacgagaaga cccccgtgag cgaccgggtg accaagtgct gcaccgagag cctggtgaac 1740

cggcggccct gcttcagcgc cctggaggtg gacgagacct acgtgcccaa ggagttcaac 1800

gccgagacct tcaccttcca cgccgacatc tgcaccctga gcgagaagga gcggcagatc 1860

aagaagcaga ccgccctggt ggagctggtg aagcacaagc ccaaggccac caaggagcag 1920

ctgaaggccg tgatggacga cttcgccgcc ttcgtggaga agtgctgcaa ggccgacgac 1980

aaggagacct gcttcgccga ggagggcaag aagctggtgg ccgccagcca ggccgccctg 2040

ggcctgggcg gcagcggcgg cagcggcggc agcggcggca gctgcagcgg ccccctgggc 2100

atcgagggcg gcatcatcag caaccagcag atcaccgcca gcagcaccca ccgggccctg 2160

ttcggcctgc agaagtggta cccctactac gcccggctga acaagaaggg cctgatcaac 2220

gcctggaccg ccgccgagaa cgaccggtgg ccctggatcc agatcaacct gcagcggaag 2280

atgcgggtga ccggcgtgat cacccagggc gccaagcgga tcggcagccc cgagtacatc 2340

aagagctaca agatcgccta cagcaacgac ggcaagacct gggccatgta caaggtgaag 2400

ggcaccaacg aggacatggt gttccggggc aacatcgaca acaacacccc ctacgccaac 2460

agcttcaccc cccccatcaa ggcccagtac gtgcggctgt acccccaggt gtgccggcgg 2520

cactgcaccc tgcggatgga gctgctgggc tgcgagctga gcggctgcag cgagcccctg 2580

ggcatgaaga gcggccacat ccaggactac cagatcaccg ccagcagcat cttccggacc 2640

ctgaacatgg acatgttcac ctgggagccc cggaaggccc ggctggacaa gcagggcaag 2700

gtgaacgcct ggaccagcgg ccacaacgac cagagccagt ggctgcaggt ggacctgctg 2760

gtgcccacca aggtgaccgg catcatcacc cagggcgcca aggacttcgg ccacgtgcag 2820

ttcgtgggca gctacaagct ggcctacagc aacgacggcg agcactggac cgtgtaccag 2880

gacgagaagc agcggaagga caaggtgttc cagggcaact tcgacaacga cacccaccgg 2940

aagaacgtga tcgacccccc catctacgcc cggcacatcc ggatcctgcc ctggagctgg 3000

tacggccgga tcaccctgcg gagcgagctg ctgggctgc 3039

<210> 94

<211> 3006

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 94

agcgccggcc cctgcacccc caacccctgc cacaacggcg gcacctgcga gatcagcgag 60

gcctaccggg gcgacacctt catcggctac gtgtgcaagt gcccccgggg cttcaacggc 120

atccactgcc agcacaacat caacgagtgc gaggtggagc cctgcaagaa cggcggcatc 180

tgcaccgacc tggtggccaa ctacagctgc gagtgccccg gcgagttcat gggccggaac 240

tgccagtaca agggcagcga cgcccacaag agcgaggtgg cccaccggtt caaggacctg 300

ggcgaggaga acttcaaggc cctggtgctg atcgccttcg cccagtacct gcagcagagc 360

cccttcgagg accacgtgaa gctggtgaac gaggtgaccg agttcgccaa gacctgcgtg 420

gccgacgaga gcgccgagaa ctgcgacaag agcctgcaca ccctgttcgg cgacaagctg 480

tgcaccgtgg ccaccctgcg ggagacctac ggcgagatgg ccgactgctg cgccaagcag 540

gagcccgagc ggaacgagtg cttcctgcag cacaaggacg acaaccccaa cctgccccgg 600

ctggtgcggc ccgaggtgga cgtgatgtgc accgccttcc acgacaacga ggagaccttc 660

ctgaagaagt acctgtacga gatcgcccgg cggcaccct acttctacgc ccccgagctg 720

ctgttcttcg ccaagcggta caaggccgcc ttcaccgagt gctgccaggc cgccgacaag 780

gccgcctgcc tgctgcccaa gctggacgag ctgcgggacg agggcaaggc cagcagcgcc 840

aagcagcggc tgaagtgcgc cagcctgcag aagttcggcg agcgggcctt caaggcctgg 900

gccgtggccc ggctgagcca gcggttcccc aaggccgagt tcgccgaggt gagcaagctg 960

gtgaccgacc tgaccaaggt gcacaccgag tgctgccacg gcgacctgct ggagtgcgcc 1020

gacgaccggg ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 1080

ctgaaggagt gctgcgagaa gcccctgctg gagaagagcc actgcatcgc cgaggtggag 1140

aacgacgaga tgcccgccga cctgcccagc ctggccgccg acttcgtgga gagcaaggac 1200

gtgtgcaaga actacgccga ggccaaggac gtgttcctgg gcatgttcct gtacgagtac 1260

gcccggcggc accccgacta cagcgtggtg ctgctgctgc ggctggccaa gacctacgag 1320

accaccctgg agaagtgctg cgccgccgcc gacccccacg agtgctacgc caaggtgttc 1380

gacgagttca agcccctggt ggaggagccc cagaacctga tcaagcagaa ctgcgagctg 1440

ttcgagcagc tgggcgagta caagttccag aacgccctgc tggtgcggta caccaagaag 1500

gtgccccagg tgagcacccc caccctggtg gaggtgagcc ggaacctggg caaggtgggc 1560

agcaagtgct gcaagcaccc cgaggccaag cggatgccct gcgccgagga ctacctgagc 1620

gtggtgctga accagctgtg cgtgctgcac gagaagaccc ccgtgagcga ccgggtgacc 1680

aagtgctgca ccgagagcct ggtgaaccgg cggccctgct tcagcgccct ggaggtggac 1740

gagacctacg tgcccaagga gttcaacgcc gagaccttca ccttccacgc cgacatctgc 1800

accctgagcg agaaggagcg gcagatcaag aagcagaccg ccctggtgga gctggtgaag 1860

cacaagccca aggccaccaa ggagcagctg aaggccgtga tggacgactt cgccgccttc 1920

gtggagaagt gctgcaaggc cgacgacaag gagacctgct tcgccgagga gggcaagaag 1980

ctggtggccg ccagccaggc cgccctgggc ctgggcggca gcggcggcag cggcggcagc 2040

ggcggcagct gcagcggccc cctgggcatc gagggcggca tcatcagcaa ccagcagatc 2100

accgccagca gcacccaccg ggccctgttc ggcctgcaga agtggtaccc ctactacgcc 2160

cggctgaaca agaagggcct gatcaacgcc tggaccgccg ccgagaacga ccggtggccc 2220

tggatccaga tcaacctgca gcggaagatg cgggtgaccg gcgtgatcac ccagggcgcc 2280

aagcggatcg gcagccccga gtacatcaag agctacaaga tcgcctacag caacgacggc 2340

aagacctggg ccatgtacaa ggtgaagggc accaacgagg acatggtgtt ccggggcaac 2400

atcgacaaca acaccccctа cgccaacagc ttcacccccc ccatcaaggc ccagtacgtg 2460

cggctgtacc cccaggtgtg ccggcggcac tgcaccctgc ggatggagct gctgggctgc 2520

gagctgagcg gctgcagcga gcccctgggc atgaagagcg gccacatcca ggactaccag 2580

atcaccgcca gcagcatctt ccggaccctg aacatggaca tgttcacctg ggagccccgg 2640

aaggcccggc tggacaagca gggcaaggtg aacgcctgga ccagcggcca caacgaccag 2700

agccagtggc tgcaggtgga cctgctggtg cccaccaagg tgaccggcat catcacccag 2760

ggcgccaagg acttcggcca cgtgcagttc gtgggcagct acaagctggc ctacagcaac 2820

gacggcgagc actggaccgt gtaccaggac gagaagcagc ggaaggacaa ggtgttccag 2880

ggcaacttcg acaacgacac ccaccggaag aacgtgatcg acccccccat ctacgcccgg 2940

cacatccgga tcctgccctg gagctggtac ggccggatca ccctgcggag cgagctgctg 3000

ggctgc 3006

<210> 95

<211> 3021

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 95

gacatctgcg accccaaccc ctgcgagaac ggcggcatct gcctgcccgg cctggccgac 60

ggcagcttca gctgcgagtg ccccgacggc ttcaccgacc ccaactgcag cagcgtggtg 120

gaggtggcca gcgacgagga ggagcccacc aacatcaacg agtgcgaggt ggagccctgc 180

aagaacggcg gcatctgcac cgacctggtg gccaactaca gctgcgagtg ccccggcgag 240

ttcatgggcc ggaactgcca gtacaagggc agcgacgccc acaagagcga ggtggcccac 300

cggttcaagg acctgggcga ggagaacttc aaggccctgg tgctgatcgc cttcgcccag 360

tacctgcagc agagcccctt cgaggaccac gtgaagctgg tgaacgaggt gaccgagttc 420

gccaagacct gcgtggccga cgagagcgcc gagaactgcg acaagagcct gcacaccctg 480

ttcggcgaca agctgtgcac cgtggccacc ctgcgggaga cctacggcga gatggccgac 540

tgctgcgcca agcaggagcc cgagcggaac gagtgcttcc tgcagcacaa ggacgacaac 600

cccaacctgc cccggctggt gcggcccgag gtggacgtga tgtgcaccgc cttccacgac 660

aacgaggaga ccttcctgaa gaagtacctg tacgagatcg cccggcggca cccctacttc 720

tacgcccccg agctgctgtt cttcgccaag cggtacaagg ccgccttcac cgagtgctgc 780

caggccgccg acaaggccgc ctgcctgctg cccaagctgg acgagctgcg ggacgagggc 840

aaggccagca gcgccaagca gcggctgaag tgcgccagcc tgcagaagtt cggcgagcgg 900

gccttcaagg cctgggccgt ggcccggctg agccagcggt tccccaaggc cgagttcgcc 960

gaggtgagca agctggtgac cgacctgacc aaggtgcaca ccgagtgctg ccacggcgac 1020

ctgctggagt gcgccgacga ccgggccgac ctggccaagt acatctgcga gaaccaggac 1080

agcatcagca gcaagctgaa ggagtgctgc gagaagcccc tgctggagaa gagccactgc 1140

atcgccgagg tggagaacga cgagatgccc gccgacctgc ccagcctggc cgccgacttc 1200

gtggagagca aggacgtgtg caagaactac gccgaggcca aggacgtgtt cctgggcatg 1260

ttcctgtacg agtacgcccg gcggcacccc gactacagcg tggtgctgct gctgcggctg 1320

gccaagacct acgagaccac cctggagaag tgctgcgccg ccgccgaccc ccacgagtgc 1380

tacgccaagg tgttcgacga gttcaagccc ctggtggagg agccccagaa cctgatcaag 1440

cagaactgcg agctgttcga gcagctgggc gagtacaagt tccagaacgc cctgctggtg 1500

cggtacacca agaaggtgcc ccaggtgagc acccccaccc tggtggaggt gagccggaac 1560

ctgggcaagg tgggcagcaa gtgctgcaag caccccgagg ccaagcggat gccctgcgcc 1620

gaggactacc tgagcgtggt gctgaaccag ctgtgcgtgc tgcacgagaa gaccccgtg 1680

agcgaccggg tgaccaagtg ctgcaccgag agcctggtga accggcggcc ctgcttcagc 1740

gccctggagg tggacgagac ctacgtgccc aaggagttca acgccgagac cttcaccttc 1800

cacgccgaca tctgcaccct gagcgagaag gagcggcaga tcaagaagca gaccgccctg 1860

gtggagctgg tgaagcacaa gcccaaggcc accaaggagc agctgaaggc cgtgatggac 1920

gacttcgccg ccttcgtgga gaagtgctgc aaggccgacg acaaggagac ctgcttcgcc 1980

gaggagggca agaagctggt ggccgccagc caggccgccc tgggcctggg cggcagcggc 2040

ggcagcggcg gcagcggcgg cagctgcagc ggccccctgg gcatcgaggg cggcatcatc 2100

agcaaccagc agatcaccgc cagcagcacc caccgggccc tgttcggcct gcagaagtgg 2160

taccccctact acgcccggct gaacaagaag ggcctgatca acgcctggac cgccgccgag 2220

aacgaccggt ggccctggat ccagatcaac ctgcagcgga agatgcgggt gaccggcgtg 2280

atcacccagg gcgccaagcg gatcggcagc cccgagtaca tcaagagcta caagatcgcc 2340

tacagcaacg acggcaagac ctgggccatg tacaaggtga agggcaccaa cgaggacatg 2400

gtgttccggg gcaacatcga caacaacacc ccctacgcca acagcttcac ccccccccatc 2460

aaggcccagt acgtgcggct gtacccccag gtgtgccggc ggcactgcac cctgcggatg 2520

gagctgctgg gctgcgagct gagcggctgc agcgagcccc tgggcatgaa gagcggccac 2580

atccaggact accagatcac cgccagcagc atcttccgga ccctgaacat ggacatgttc 2640

acctgggagc cccggaaggc ccggctggac aagcagggca aggtgaacgc ctggaccagc 2700

ggccacaacg accagagcca gtggctgcag gtggacctgc tggtgcccac caaggtgacc 2760

ggcatcatca cccagggcgc caaggacttc ggccacgtgc agttcgtggg cagctacaag 2820

ctggcctaca gcaacgacgg cgagcactgg accgtgtacc aggacgagaa gcagcggaag 2880

gacaaggtgt tccagggcaa cttcgacaac gacacccacc ggaagaacgt gatcgacccc 2940

cccatctacg cccggcacat ccggatcctg ccctggagct ggtacggccg gatcaccctg 3000

cggagcgagc tgctgggctg c 3021

<210> 96

<211> 50

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 96

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro

1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr

20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu

35 40 45

Pro Thr

50

<210> 97

<211> 45

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 97

Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly Thr Cys

1 5 10 15

Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr Val Cys

20 25 30

Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His

35 40 45

<210> 98

<211> 39

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 98

Asn Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys

1 5 10 15

Thr Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met

20 25 30

Gly Arg Asn Cys Gln Tyr Lys

35

<210> 99

<211> 95

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 99

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro

1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr

20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu

35 40 45

Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly

50 55 60

Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr
65 70 75 80
Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His
85 90 95

<210> 100
<211> 84
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 100
Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly Thr Cys
1 5 10 15
Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr Val Cys
20 25 30
Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn Ile Asn
35 40 45
Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr Asp Leu
50 55 60
Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly Arg Asn
65 70 75 80
Cys Gln Tyr Lys

<210> 101
<211> 89
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 101

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15
Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30
Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45
Pro Thr Asn Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly
50 55 60
Ile Cys Thr Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu
65 70 75 80
Phe Met Gly Arg Asn Cys Gln Tyr Lys
85

<210> 102

<211> 945

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 102

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15
Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30
Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45
Pro Thr Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu
50 55 60
Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr
65 70 75 80

Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val
85 90 95
Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys
100 105 110
Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala
115 120 125
Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln
130 135 140
Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro
145 150 155 160
Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala
165 170 175
Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile
180 185 190
Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala
195 200 205
Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys
210 215 220
Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys
225 230 235 240
Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe
245 250 255
Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg
260 265 270
Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu
275 280 285
Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala
290 295 300
Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser
305 310 315 320
Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys

325 330 335

Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu

340 345 350

Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn

355 360 365

Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr

370 375 380

Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala

385 390 395 400

Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro

405 410 415

His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu

420 425 430

Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu

435 440 445

Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys

450 455 460

Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu

465 470 475 480

Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met

485 490 495

Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val

500 505 510

Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr

515 520 525

Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp

530 535 540

Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His

545 550 555 560

Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln

565 570 575

Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu
580 585 590
Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys
595 600 605
Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys
610 615 620
Leu Val Ala Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala
625 630 635 640
Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu
645 650 655
Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val
660 665 670
Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val
675 680 685
Asn Leu Leu Arg Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala
690 695 700
Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr
705 710 715 720
Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys
725 730 735
His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn
740 745 750
Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr
755 760 765
Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu
770 775 780
Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro
785 790 795 800
Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His
805 810 815
Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn

820 825 830

Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln

835 840 845

Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly

850 855 860

Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala

865 870 875 880

Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr

885 890 895

Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys

900 905 910

Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu

915 920 925

Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly

930 935 940

Cys

945

<210> 103

<211> 940

<212> PRT

<213> Trình tự nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 103

Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly Thr Cys

1 5 10 15

Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr Val Cys

20 25 30

Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asp Ala His

35 40 45

Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe
50 55 60
Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser Pro
65 70 75 80
Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys
85 90 95
Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His
100 105 110
Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr
115 120 125
Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn
130 135 140
Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu
145 150 155 160
Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn Glu
165 170 175
Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro
180 185 190
Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala
195 200 205
Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu
210 215 220
Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys
225 230 235 240
Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe
245 250 255
Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu
260 265 270
Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His Thr
275 280 285
Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp

290 295 300

Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu

305 310 315 320

Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile Ala

325 330 335

Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala Ala

340 345 350

Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys

355 360 365

Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro

370 375 380

Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr

385 390 395 400

Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr Ala

405 410 415

Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn Leu

420 425 430

Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe

435 440 445

Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val Ser

450 455 460

Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly Ser

465 470 475 480

Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp

485 490 495

Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr

500 505 510

Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn

515 520 525

Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro

530 535 540

Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr
545 550 555 560
Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu
565 570 575
Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val
580 585 590
Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp
595 600 605
Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Cys Val
610 615 620
Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile Ala
625 630 635 640
Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val Pro
645 650 655
Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr Pro
660 665 670
Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg Arg
675 680 685
Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala Ser
690 695 700
His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly His
705 710 715 720
Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe Val
725 730 735
Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr Pro
740 745 750
Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr Ala
755 760 765
Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys Ala
770 775 780
Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile Thr

785 790 795 800

Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp Asn

805 810 815

Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp Val

820 825 830

Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly Ser

835 840 845

Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly

850 855 860

Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser

865 870 875 880

Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys Ile

885 890 895

Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe Glu

900 905 910

Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp His

915 920 925

Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys

930 935 940

<210> 104

<211> 934

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 104

Asn Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys

1 5 10 15

Thr Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met

20 25 30

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

275 280 285

Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys

290 295 300

Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys

305 310 315 320

Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu

325 330 335

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

340 345 350

Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met

355 360 365

Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu

370 375 380

Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys

385 390 395 400

Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe

405 410 415

Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu

420 425 430

Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val

435 440 445

Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu

450 455 460

Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro

465 470 475 480

Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu

485 490 495

Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val

500 505 510

Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser

515 520 525

Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu
530 535 540
Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg
545 550 555 560
Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro
565 570 575
Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala
580 585 590
Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala
595 600 605
Glu Glu Gly Lys Lys Leu Val Ala Cys Val Glu Pro Leu Gly Met Glu
610 615 620
Asn Gly Asn Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val
625 630 635 640
Thr Phe Leu Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn
645 650 655
Arg Ala Gly Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn
660 665 670
Pro Trp Ile Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr Gly Val
675 680 685
Val Thr Gln Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala
690 695 700
Phe Lys Val Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His
705 710 715 720
Asp Val Asn Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn
725 730 735
Ala Val His Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val
740 745 750
Arg Leu Tyr Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu
755 760 765
Leu Leu Gly Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys

770 775 780

Asn Asn Ser Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys
785 790 795 800

Thr Trp Gly Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu
805 810 815

Asp Lys Gln Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn
820 825 830

Asp Gln Trp Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly
835 840 845

Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala
850 855 860

Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr
865 870 875 880

Gln Asp Pro Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp
825 890 895

Asn His Ser His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg
900 905 910

Tyr Val Arg Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg
915 920 925

Leu Glu Leu Leu Gly Cys
930

<210> 105
<211> 990
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 105

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30
Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45
Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly
50 55 60
Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr
65 70 75 80
Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asp
85 90 95
Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu
100 105 110
Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln
115 120 125
Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe
130 135 140
Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser
145 150 155 160
Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg
165 170 175
Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu
180 185 190
Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro
195 200 205
Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp
210 215 220
Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg
225 230 235 240
His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr
245 250 255
Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys

260 265 270

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

275 280 285

Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg

290 295 300

Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys

305 310 315 320

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

325 330 335

His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg

340 345 350

Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser

355 360 365

Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys

370 375 380

Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu

385 390 395 400

Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu

405 410 415

Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg

420 425 430

His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr

435 440 445

Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys

450 455 460

Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln

465 470 475 480

Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr

485 490 495

Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln

500 505 510

Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val
515 520 525

Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala
530 535 540

Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu
545 550 555 560

Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu
565 570 575

Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr
580 585 590

Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile
595 600 605

Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu
610 615 620

Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys
625 630 635 640

Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala
645 650 655

Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala
660 665 670

Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln
675 680 685

Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp
690 695 700

Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp
705 710 715 720

Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu
725 730 735

Arg Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu
740 745 750

Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn

755 760 765

Gly His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu

770 775 780

Phe Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu

785 790 795 800

Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His

805 810 815

Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly

820 825 830

Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln

835 840 845

Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser

850 855 860

Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala

865 870 875 880

Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu

885 890 895

Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn

900 905 910

Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn

915 920 925

Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser

930 935 940

Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu

945 950 955 960

Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala

965 970 975

Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys

980 985 990

<210> 106

<211> 979

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 106

Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly Thr Cys
1 5 10 15

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

Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn Ile Asn
35 40 45

Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr Asp Leu
50 55 60

Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly Arg Asn
65 70 75 80

Cys Gln Tyr Lys Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys
85 90 95

Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala
100 105 110

Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn
115 120 125

Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu
130 135 140

Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr
145 150 155 160

Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala
165 170 175

Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp
180 185 190

Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys

195 200 205

Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr

210 215 220

Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe

225 230 235 240

Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala

245 250 255

Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu

260 265 270

Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln

275 280 285

Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser

290 295 300

Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr

305 310 315 320

Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu

325 330 335

Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln

340 345 350

Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu

355 360 365

Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala

370 375 380

Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys

385 390 395 400

Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr

405 410 415

Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg

420 425 430

Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala

435 440 445

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

690 695 700

Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile

705 710 715 720

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

725 730 735

Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val

740 745 750

Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn

755 760 765

Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His

770 775 780

Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr

785 790 795 800

Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly

805 810 815

Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser

820 825 830

Ile Pro Asp Lys Gln Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly

835 840 845

Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln

850 855 860

Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp

865 870 875 880

Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr

885 890 895

Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys

900 905 910

Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro

915 920 925

Arg Thr Gly Ser Ser Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser

930 935 940

His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg
945 950 955 960
Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu
965 970 975
Leu Gly Cys

<210> 107
<211> 984
<212> PRT
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"
<400> 107
Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15
Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30
Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45
Pro Thr Asn Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly
50 55 60
Ile Cys Thr Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu
65 70 75 80
Phe Met Gly Arg Asn Cys Gln Tyr Lys Asp Ala His Lys Ser Glu Val
85 90 95
Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val
100 105 110
Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His
115 120 125
Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala

130 135 140

Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly

145 150 155 160

Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met

165 170 175

Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu

180 185 190

Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu

195 200 205

Val Asp Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu

210 215 220

Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala

225 230 235 240

Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu

245 250 255

Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp

260 265 270

Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys

275 280 285

Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala

290 295 300

Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val

305 310 315 320

Ser Lys Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His

325 330 335

Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr

340 345 350

Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys

355 360 365

Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn

370 375 380

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

625 630 635 640

Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys

645 650 655

Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Cys Val Glu Pro Leu Gly

660 665 670

Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val

675 680 685

Arg Val Thr Phe Leu Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg

690 695 700

Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp

705 710 715 720

Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr

725 730 735

Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu

740 745 750

Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe

755 760 765

Ile His Asp Val Asn Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn

770 775 780

Lys Asn Ala Val His Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln

785 790 795 800

Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg

805 810 815

Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys Ala Asn Pro Leu Gly

820 825 830

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

835 840 845

Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp Asn Pro Ser Tyr Ala

850 855 860

Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp Val Ala Gly Ser Tyr

865 870 875 880

Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly Ser Ser Lys Glu Val
885 890 895

Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe Gly Ser Val Gln Phe
900 905 910

Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp Ser Ala Asn Trp Thr
915 920 925

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

Trp Asp Asn His Ser His Lys Lys Asn Leu Phe Glu Thr Pro Ile Leu
945 950 955 960

Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp His Asn Arg Ile Ala
965 970 975

Leu Arg Leu Glu Leu Leu Gly Cys
980

<210> 108

<211> 2835

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 108

gacatctgcg accccaaccc ctgcgagaac ggcggcatct gcctgcccgg cctggccgac 60

ggcagcttca gctgcgagtg ccccgacggc ttcaccgacc ccaactgcag cagcgtggtg 120

gaggtggcca gcgacgagga ggagcccacc gacgcccaca agagcgaggt ggcccaccgg 180

ttcaaggacc tgggcgagga gaacttcaag gccctggtgc tgatcgcctt cgcccagtac 240

ctgcagcaga gcccccttcga ggaccacgtg aagctggtga acgaggtgac cgagttcgcc 300

aagacctgcg tggccgacga gagcgccgag aactgcgaca agagcctgca caccctgttc 360

ggcgacaagc tgtgcaccgt ggccaccctg cgggagacct acggcgagat ggccgactgc 420

tgcgccaagc aggagcccga gcggaacgag tgcttcctgc agcacaagga cgacaacccc 480

aacctgcccc ggctggtgcg gcccgaggtg gacgtgatgt gcaccgcctt ccacgacaac 540

gaggagacct tcctgaagaa gtacctgtac gagatcgccc ggcggcaccc ctacttctac 600

gcccccgagc tgctgttctt cgccaagcgg tacaaggccg ccttcaccga gtgctgccag 660

gccgccgaca aggccgcctg cctgctgccc aagctggacg agctgcggga cgagggcaag 720

gccagcagcg ccaagcagcg gctgaagtgc gccagcctgc agaagttcgg cgagcgggcc 780

ttcaaggcct gggccgtggc ccggctgagc cagcggttcc ccaaggccga gttcgccgag 840

gtgagcaagc tggtgaccga cctgaccaag gtgcacaccg agtgctgcca cggcgacctg 900

ctggagtgcg ccgacgaccg ggccgacctg gccaagtaca tctgcgagaa ccaggacagc 960

atcagcagca agctgaagga gtgctgcgag aagcccctgc tggagaagag ccactgcatc 1020

gccgaggtgg agaacgacga gatgcccgcc gacctgccca gcctggccgc cgacttcgtg 1080

gagagcaagg acgtgtgcaa gaactacgcc gaggccaagg acgtgttcct gggcatgttc 1140

ctgtacgagt acgcccggcg gcaccccgac tacagcgtgg tgctgctgct gcggctggcc 1200

aagacctacg agaccaccct ggagaagtgc tgcgccgccg ccgaccccca cgagtgctac 1260

gccaaggtgt tcgacgagtt caagcccctg gtggaggagc cccagaacct gatcaagcag 1320

aactgcgagc tgttcgagca gctgggcgag tacaagttcc agaacgccct gctggtgcgg 1380

tacaccaaga aggtgcccca ggtgagcacc cccaccctgg tggaggtgag ccggaacctg 1440

ggcaaggtgg gcagcaagtg ctgcaagcac cccgaggcca agcggatgcc ctgcgccgag 1500

gactacctga gcgtggtgct gaaccagctg tgcgtgctgc acgagaagac ccccgtgagc 1560

gaccgggtga ccaagtgctg caccgagagc ctggtgaacc ggcggccctg cttcagcgcc 1620

ctggaggtgg acgagaccta cgtgcccaag gagttcaacg ccgagacctt caccttccac 1680

gccgacatct gcaccctgag cgagaaggag cggcagatca agaagcagac cgccctggtg 1740

gagctggtga agcacaagcc caaggccacc aaggagcagc tgaaggccgt gatggacgac 1800

ttcgccgcct tcgtggagaa gtgctgcaag gccgacgaca aggagacctg cttcgccgag 1860

gagggcaaga agctggtggc ctgcgtggag cccctgggca tggagaacgg caacatcgcc 1920

aacagccaga tcgccgccag cagcgtgcgg gtgaccttcc tgggcctgca gcactgggtg 1980

cccgagctgg cccggctgaa ccgggccggc atggtgaacg cctggacccc cagcagcaac 2040

gacgacaacc cctggatcca ggtgaacctg ctgcggcgga tgtgggtgac cggcgtggtg 2100

acccagggcg ccagccggct ggccagccac gagtacctga aggccttcaa ggtggcctac 2160

agcctgaacg gccacgagtt cgacttcatc cacgacgtga acaagaagca caaggagttc 2220

gtgggcaact ggaacaagaa cgccgtgcac gtgaacctgt tcgagacccc cgtggaggcc 2280

cagtacgtgc ggctgtaccc caccagctgc cacaccgcct gcaccctgcg gttcgagctg 2340

ctgggctgcg agctgaacgg ctgcgccaac cccctgggcc tgaagaacaa cagcatcccc 2400

gacaagcaga tcaccgccag cagcagctac aagacctggg gcctgcacct gttcagctgg 2460

aaccccagct acgcccggct ggacaagcag ggcaacttca acgcctgggt ggccggcagc 2520

tacggcaacg accagtggct gcaggtggac ctgggcagca gcaaggaggt gaccggcatc 2580

atcacccagg gcgcccggaa cttcggcagc gtgcagttcg tggccagcta caaggtggcc 2640

tacagcaacg acagcgccaa ctggaccgag taccaggacc cccggaccgg cagcagcaag 2700

atcttccccg gcaactggga caaccacagc cacaagaaga acctgttcga gacccccatc 2760

ctggcccggt acgtgcggat cctgcccgtg gcctggcaca accggatcgc cctgcggctg 2820

gagctgctgg gctgc 2835

<210> 109

<211> 2820

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 109

agcgccggcc cctgcacccc caacccctgc cacaacggcg gcacctgcga gatcagcgag 60

gcctaccggg gcgacacctt catcggctac gtgtgcaagt gcccccgggg cttcaacggc 120

atccactgcc agcacgacgc ccacaagagc gaggtggccc accggttcaa ggacctgggc 180

gaggagaact tcaaggccct ggtgctgatc gccttcgccc agtacctgca gcagagcccc 240

ttcgaggacc acgtgaagct ggtgaacgag gtgaccgagt tcgccaagac ctgcgtggcc 300

gacgagagcg ccgagaactg cgacaagagc ctgcacaccc tgttcggcga caagctgtgc 360

accgtggcca ccctgcggga gacctacggc gagatggccg actgctgcgc caagcaggag 420

cccgagcgga acgagtgctt cctgcagcac aaggacgaca accccaacct gccccggctg 480

gtgcggcccg aggtggacgt gatgtgcacc gccttccacg acaacgagga gaccttcctg 540

aagaagtacc tgtacgagat cgcccggcgg cacccctact tctacgcccc cgagctgctg 600

ttcttcgcca agcggtacaa ggccgccttc accgagtgct gccaggccgc cgacaaggcc 660

gcctgcctgc tgcccaagct ggacgagctg cgggacgagg gcaaggccag cagcgccaag 720

cagcggctga agtgcgccag cctgcagaag ttcggcgagc gggccttcaa ggcctgggcc 780

gtggcccggc tgagccagcg gttccccaag gccgagttcg ccgaggtgag caagctggtg 840

accgacctga ccaaggtgca caccgagtgc tgccacggcg acctgctgga gtgcgccgac 900

gaccgggccg acctggccaa gtacatctgc gagaaccagg acagcatcag cagcaagctg 960

aaggagtgct gcgagaagcc cctgctggag aagagccact gcatcgccga ggtggagaac 1020

gacgagatgc ccgccgacct gcccagcctg gccgccgact tcgtggagag caaggacgtg 1080

tgcaagaact acgccgaggc caaggacgtg ttcctgggca tgttcctgta cgagtacgcc 1140

cggcggcacc ccgactacag cgtggtgctg ctgctgcggc tggccaagac ctacgagacc 1200

accctggaga agtgctgcgc cgccgccgac ccccacgagt gctacgccaa ggtgttcgac 1260

gagttcaagc ccctggtgga ggagccccag aacctgatca agcagaactg cgagctgttc 1320

gagcagctgg gcgagtacaa gttccagaac gccctgctgg tgcggtacac caagaaggtg 1380

ccccaggtga gcacccccac cctggtggag gtgagccgga acctgggcaa ggtgggcagc 1440

aagtgctgca agcaccccga ggccaagcgg atgccctgcg ccgaggacta cctgagcgtg 1500

gtgctgaacc agctgtgcgt gctgcacgag aagacccccg tgagcgaccg ggtgaccaag 1560

tgctgcaccg agagcctggt gaaccggcgg ccctgcttca gcgccctgga ggtggacgag 1620

acctacgtgc ccaaggagtt caacgccgag accttcacct tccacgccga catctgcacc 1680

ctgagcgaga aggagcggca gatcaagaag cagaccgccc tggtggagct ggtgaagcac 1740

aagcccaagg ccaccaagga gcagctgaag gccgtgatgg acgacttcgc cgccttcgtg 1800

gagaagtgct gcaaggccga cgacaaggag acctgcttcg ccgaggaggg caagaagctg 1860

gtggcctgcg tggagcccct gggcatggag aacggcaaca tcgccaacag ccagatcgcc 1920

gccagcagcg tgcgggtgac cttcctgggc ctgcagcact gggtgcccga gctggcccgg 1980

ctgaaccggg ccggcatggt gaacgcctgg acccccagca gcaacgacga caacccctgg 2040

atccaggtga acctgctgcg gcggatgtgg gtgaccggcg tggtgaccca gggcgccagc 2100

cggctggcca gccacgagta cctgaaggcc ttcaaggtgg cctacagcct gaacggccac 2160

gagttcgact tcatccacga cgtgaacaag aagcacaagg agttcgtggg caactggaac 2220

aagaacgccg tgcacgtgaa cctgttcgag acccccgtgg aggcccagta cgtgcggctg 2280

taccccacca gctgccacac cgcctgcacc ctgcggttcg agctgctggg ctgcgagctg 2340

aacggctgcg ccaacccccc gggcctgaag aacaacagca tccccgacaa gcagatcacc 2400

gccagcagca gctacaagac ctggggcctg cacctgttca gctggaaccc cagctacgcc 2460

cggctggaca agcagggcaa cttcaacgcc tgggtggccg gcagctacgg caacgaccag 2520

tggctgcagg tggacctggg cagcagcaag gaggtgaccg gcatcatcac ccagggcgcc 2580

cggaacttcg gcagcgtgca gttcgtggcc agctacaagg tggcctacag caacgacagc 2640

gccaactgga ccgagtacca ggacccccgg accggcagca gcaagatctt ccccggcaac 2700

tgggacaacc acagccacaa gaagaacctg ttcgagaccc ccatcctggc ccggtacgtg 2760

cggatcctgc ccgtggcctg gcacaaccgg atcgccctgc ggctggagct gctgggctgc 2820

<210> 110

<211> 2802

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 110

aacatcaacg agtgcgaggt ggagccctgc aagaacggcg gcatctgcac cgacctggtg 60

gccaactaca gctgcgagtg ccccggcgag ttcatgggcc ggaactgcca gtacaaggac 120

gcccacaaga gcgaggtggc ccaccggttc aaggacctgg gcgaggagaa cttcaaggcc 180

ctggtgctga tcgccttcgc ccagtacctg cagcagagcc ccttcgagga ccacgtgaag 240

ctggtgaacg aggtgaccga gttcgccaag acctgcgtgg ccgacgagag cgccgagaac 300

tgcgacaaga gcctgcacac cctgttcggc gacaagctgt gcaccgtggc caccctgcgg 360

gagacctacg gcgagatggc cgactgctgc gccaagcagg agcccgagcg gaacgagtgc 420

ttcctgcagc acaaggacga caaccccaac ctgccccggc tggtgcggcc cgaggtggac 480

gtgatgtgca ccgccttcca cgacaacgag gagaccttcc tgaagaagta cctgtacgag 540

atcgcccggc ggcaccccta cttctacgcc cccgagctgc tgttcttcgc caagcggtac 600

aaggccgcct tcaccgagtg ctgccaggcc gccgacaagg ccgcctgcct gctgcccaag 660

ctggacgagc tgcgggacga gggcaaggcc agcagcgcca agcagcggct gaagtgcgcc 720

agcctgcaga agttcggcga gcgggccttc aaggcctggg ccgtggcccg gctgagccag 780

cggttcccca aggccgagtt cgccgaggtg agcaagctgg tgaccgacct gaccaaggtg 840

cacaccgagt gctgccacgg cgacctgctg gagtgcgccg acgaccgggc cgacctggcc 900

aagtacatct gcgagaacca ggacagcatc agcagcaagc tgaaggagtg ctgcgagaag 960

cccctgctgg agaagagcca ctgcatcgcc gaggtggaga acgacgagat gcccgccgac 1020

ctgcccagcc tggccgccga cttcgtggag agcaaggacg tgtgcaagaa ctacgccgag 1080

gccaaggacg tgttcctggg catgttcctg tacgagtacg cccggcggca ccccgactac 1140

agcgtggtgc tgctgctgcg gctggccaag acctacgaga ccaccctgga gaagtgctgc 1200

gccgccgccg acccccacga gtgctacgcc aaggtgttcg acgagttcaa gcccctggtg 1260

gaggagcccc agaacctgat caagcagaac tgcgagctgt tcgagcagct gggcgagtac 1320

aagttccaga acgccctgct ggtgcggtac accaagaagg tgccccaggt gagcacccc 1380

accctggtgg aggtgagccg gaacctgggc aaggtgggca gcaagtgctg caagcacccc 1440

gaggccaagc ggatgccctg cgccgaggac tacctgagcg tggtgctgaa ccagctgtgc 1500

gtgctgcacg agaagacccc cgtgagcgac cgggtgacca agtgctgcac cgagagcctg 1560

gtgaaccggc ggccctgctt cagcgccctg gaggtggacg agacctacgt gcccaaggag 1620

ttcaacgccg agaccttcac cttccacgcc gacatctgca ccctgagcga gaaggagcgg 1680

cagatcaaga agcagaccgc cctggtggag ctggtgaagc acaagcccaa ggccaccaag 1740

gagcagctga aggccgtgat ggacgacttc gccgccttcg tggagaagtg ctgcaaggcc 1800

gacgacaagg agacctgctt cgccgaggag ggcaagaagc tggtggcctg cgtggagccc 1860

ctgggcatgg agaacggcaa catcgccaac agccagatcg ccgccagcag cgtgcgggtg 1920

accttcctgg gcctgcagca ctgggtgccc gagctggccc ggctgaaccg ggccggcatg 1980

gtgaacgcct ggacccccag cagcaacgac gacaacccct ggatccaggt gaacctgctg 2040

cggcggatgt gggtgaccgg cgtggtgacc cagggcgcca gccggctggc cagccacgag 2100

tacctgaagg ccttcaaggt ggcctacagc ctgaacggcc acgagttcga cttcatccac 2160

gacgtgaaca agaagcacaa ggagttcgtg ggcaactgga acaagaacgc cgtgcacgtg 2220

aacctgttcg agacccccgt ggaggcccag tacgtgcggc tgtaccccac cagctgccac 2280

accgcctgca ccctgcggtt cgagctgctg ggctgcgagc tgaacggctg cgccaacccc 2340

ctgggcctga agaacaacag catccccgac aagcagatca ccgccagcag cagctacaag 2400

acctggggcc tgcacctgtt cagctggaac cccagctacg cccggctgga caagcagggc 2460

aacttcaacg cctgggtggc cggcagctac ggcaacgacc agtggctgca ggtggacctg 2520

ggcagcagca aggaggtgac cggcatcatc acccagggcg cccggaactt cggcagcgtg 2580

cagttcgtgg ccagctacaa ggtggcctac agcaacgaca gcgccaactg gaccgagtac 2640

caggaccccc ggaccggcag cagcaagatc ttccccggca actgggacaa ccacagccac 2700

aagaagaacc tgttcgagac ccccatcctg gcccggtacg tgcggatcct gcccgtggcc 2760

tggcacaacc ggatcgccct gcggctggag ctgctgggct gc 2802

<210> 111

<211> 2970

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 111

gacatctgcg accccaaccc ctgcgagaac ggcggcatct gcctgcccgg cctggccgac 60

ggcagcttca gctgcgagtg ccccgacggc ttcaccgacc ccaactgcag cagcgtggtg 120

gaggtggcca gcgacgagga ggagcccacc agcgccggcc cctgcacccc caacccctgc 180

cacaacggcg gcacctgcga gatcagcgag gcctaccggg gcgacacctt catcggctac 240

gtgtgcaagt gcccccgggg cttcaacggc atccactgcc agcacgacgc ccacaagagc 300

gaggtggccc accggttcaa ggacctgggc gaggagaact tcaaggccct ggtgctgatc 360

gccttcgccc agtacctgca gcagagcccc ttcgaggacc acgtgaagct ggtgaacgag 420

gtgaccgagt tcgccaagac ctgcgtggcc gacgagagcg ccgagaactg cgacaagagc 480

ctgcacaccc tgttcggcga caagctgtgc accgtggcca ccctgcggga gacctacggc 540

gagatggccg actgctgcgc caagcaggag cccgagcgga acgagtgctt cctgcagcac 600

aaggacgaca accccaacct gccccggctg gtgcggcccg aggtggacgt gatgtgcacc 660

gccttccacg acaacgagga gaccttcctg aagaagtacc tgtacgagat cgcccggcgg 720

cacccctact tctacgcccc cgagctgctg ttcttcgcca agcggtacaa ggccgccttc 780

accgagtgct gccaggccgc cgacaaggcc gcctgcctgc tgcccaagct ggacgagctg 840

cgggacgagg gcaaggccag cagcgccaag cagcggctga agtgcgccag cctgcagaag 900

ttcggcgagc gggccttcaa ggcctgggcc gtggcccggc tgagccagcg gttccccaag 960

gccgagttcg ccgaggtgag caagctggtg accgacctga ccaaggtgca caccgagtgc 1020

tgccacggcg acctgctgga gtgcgccgac gaccgggccg acctggccaa gtacatctgc 1080

gagaaccagg acagcatcag cagcaagctg aaggagtgct gcgagaagcc cctgctggag 1140

aagagccact gcatcgccga ggtggagaac gacgagatgc ccgccgacct gcccagcctg 1200

gccgccgact tcgtggagag caaggacgtg tgcaagaact acgccgaggc caaggacgtg 1260

ttcctgggca tgttcctgta cgagtacgcc cggcggcacc ccgactacag cgtggtgctg 1320

ctgctgcggc tggccaagac ctacgagacc accctggaga agtgctgcgc cgccgccgac 1380

ccccacgagt gctacgccaa ggtgttcgac gagttcaagc ccctggtgga ggagccccag 1440

aacctgatca agcagaactg cgagctgttc gagcagctgg gcgagtacaa gttccagaac 1500

gccctgctgg tgcggtacac caagaaggtg ccccaggtga gcacccccac cctggtggag 1560

gtgagccgga acctgggcaa ggtgggcagc aagtgctgca agcaccccga ggccaagcgg 1620

atgccctgcg ccgaggacta cctgagcgtg gtgctgaacc agctgtgcgt gctgcacgag 1680

aagacccccg tgagcgaccg ggtgaccaag tgctgcaccg agagcctggt gaaccggcgg 1740

ccctgcttca gcgccctgga ggtggacgag acctacgtgc ccaaggagtt caacgccgag 1800

accttcacct tccacgccga catctgcacc ctgagcgaga aggagcggca gatcaagaag 1860

cagaccgccc tggtggagct ggtgaagcac aagcccaagg ccaccaagga gcagctgaag 1920

gccgtgatgg acgacttcgc cgccttcgtg gagaagtgct gcaaggccga cgacaaggag 1980

acctgcttcg ccgaggaggg caagaagctg gtggcctgcg tggagcccct gggcatggag 2040

aacggcaaca tcgccaacag ccagatcgcc gccagcagcg tgcgggtgac cttcctgggc 2100

ctgcagcact gggtgcccga gctggcccgg ctgaaccggg ccggcatggt gaacgcctgg 2160

acccccagca gcaacgacga caacccctgg atccaggtga acctgctgcg gcggatgtgg 2220

gtgaccggcg tggtgaccca gggcgccagc cggctggcca gccacgagta cctgaaggcc 2280

ttcaaggtgg cctacagcct gaacggccac gagttcgact tcatccacga cgtgaacaag 2340

aagcacaagg agttcgtggg caactggaac aagaacgccg tgcacgtgaa cctgttcgag 2400

acccccgtgg aggcccagta cgtgcggctg taccccacca gctgccacac cgcctgcacc 2460

ctgcggttcg agctgctggg ctgcgagctg aacggctgcg ccaacccccc gggcctgaag 2520

aacaacagca tccccgacaa gcagatcacc gccagcagca gctacaagac ctggggcctg 2580

cacctgttca gctggaaccc cagctacgcc cggctggaca agcagggcaa cttcaacgcc 2640

tgggtggccg gcagctacgg caacgaccag tggctgcagg tggacctggg cagcagcaag 2700

gaggtgaccg gcatcatcac ccagggcgcc cggaacttcg gcagcgtgca gttcgtggcc 2760

agctacaagg tggcctacag caacgacagc gccaactgga ccgagtacca ggacccccgg 2820

accggcagca gcaagatctt ccccggcaac tgggacaacc acagccacaa gaagaacctg 2880

ttcgagaccc ccatcctggc ccggtacgtg cggatcctgc ccgtggcctg gcacaaccgg 2940

atcgccctgc ggctggagct gctgggctgc 2970

<210> 112

<211> 2937

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 112

agcgccggcc cctgcacccc caacccctgc cacaacggcg gcacctgcga gatcagcgag 60

gcctaccggg gcgacacctt catcggctac gtgtgcaagt gccccccggg cttcaacggc 120

atccactgcc agcacaacat caacgagtgc gaggtggagc cctgcaagaa cggcggcatc 180

tgcaccgacc tggtggccaa ctacagctgc gagtgccccg gcgagttcat gggccggaac 240

tgccagtaca aggacgccca caagagcgag gtggcccacc ggttcaagga cctgggcgag 300

gagaacttca aggccctggt gctgatcgcc ttcgcccagt acctgcagca gagccccttc 360

gaggaccacg tgaagctggt gaacgaggtg accgagttcg ccaagacctg cgtggccgac 420

gagagcgccg agaactgcga caagagcctg cacaccctgt tcggcgacaa gctgtgcacc 480

gtggccaccc tgcgggagac ctacggcgag atggccgact gctgcgccaa gcaggagccc 540

gagcggaacg agtgcttcct gcagcacaag gacgacaacc ccaacctgcc ccggctggtg 600

cggcccgagg tggacgtgat gtgcaccgcc ttccacgaca acgaggagac cttcctgaag 660

aagtacctgt acgagatcgc ccggcggcac ccctacttct acgcccccga gctgctgttc 720

ttcgccaagc ggtacaaggc cgccttcacc gagtgctgcc aggccgccga caaggccgcc 780

tgcctgctgc ccaagctgga cgagctgcgg gacgagggca aggccagcag cgccaagcag 840

cggctgaagt gcgccagcct gcagaagttc ggcgagcggg ccttcaaggc ctgggccgtg 900

gcccggctga gccagcggtt ccccaaggcc gagttcgccg aggtgagcaa gctggtgacc 960

gacctgacca aggtgcacac cgagtgctgc cacggcgacc tgctggagtg cgccgacgac 1020

cgggccgacc tggccaagta catctgcgag aaccaggaca gcatcagcag caagctgaag 1080

gagtgctgcg agaagcccct gctggagaag agccactgca tcgccgaggt ggagaacgac 1140

gagatgcccg ccgacctgcc cagcctggcc gccgacttcg tggagagcaa ggacgtgtgc 1200

aagaactacg ccgaggccaa ggacgtgttc ctgggcatgt tcctgtacga gtacgcccgg 1260

cggcaccccg actacagcgt ggtgctgctg ctgcggctgg ccaagaccta cgagaccacc 1320

ctggagaagt gctgcgccgc cgccgacccc cacgagtgct acgccaaggt gttcgacgag 1380

ttcaagcccc tggtggagga gccccagaac ctgatcaagc agaactgcga gctgttcgag 1440

cagctgggcg agtacaagtt ccagaacgcc ctgctggtgc ggtacaccaa gaaggtgccc 1500

caggtgagca cccccaccct ggtggaggtg agccggaacc tgggcaaggt gggcagcaag 1560

tgctgcaagc accccgaggc caagcggatg ccctgcgccg aggactacct gagcgtggtg 1620

ctgaaccagc tgtgcgtgct gcacgagaag acccccgtga gcgaccgggt gaccaagtgc 1680

tgcaccgaga gcctggtgaa ccggcggccc tgcttcagcg ccctggaggt ggacgagacc 1740

tacgtgccca aggagttcaa cgccgagacc ttcaccttcc acgccgacat ctgcaccctg 1800

agcgagaagg agcggcagat caagaagcag accgccctgg tggagctggt gaagcacaag 1860

cccaaggcca ccaaggagca gctgaaggcc gtgatggacg acttcgccgc cttcgtggag 1920

aagtgctgca aggccgacga caaggagacc tgcttcgccg aggagggcaa gaagctggtg 1980

gcctgcgtgg agcccctggg catggagaac ggcaacatcg ccaacagcca gatcgccgcc 2040

agcagcgtgc gggtgacctt cctgggcctg cagcactggg tgcccgagct ggcccggctg 2100

aaccgggccg gcatggtgaa cgcctggacc cccagcagca acgacgacaa cccctggatc 2160

caggtgaacc tgctgcggcg gatgtgggtg accggcgtgg tgacccaggg cgccagccgg 2220

ctggccagcc acgagtacct gaaggccttc aaggtggcct acagcctgaa cggccacgag 2280

ttcgacttca tccacgacgt gaacaagaag cacaaggagt tcgtgggcaa ctggaacaag 2340

aacgccgtgc acgtgaacct gttcgagacc cccgtggagg cccagtacgt gcggctgtac 2400

cccaccagct gccacaccgc ctgcaccctg cggttcgagc tgctgggctg cgagctgaac 2460

ggctgcgcca acccccctggg cctgaagaac aacagcatcc ccgacaagca gatcaccgcc 2520

agcagcagct acaagacctg gggcctgcac ctgttcagct ggaaccccag ctacgcccgg 2580

ctggacaagc agggcaactt caacgcctgg gtggccggca gctacggcaa cgaccagtgg 2640

ctgcaggtgg acctgggcag cagcaaggag gtgaccggca tcatcaccca gggcgcccgg 2700

aacttcggca gcgtgcagtt cgtggccagc tacaaggtgg cctacagcaa cgacagcgcc 2760

aactggaccg agtaccagga cccccggacc ggcagcagca agatcttccc cggcaactgg 2820

gacaaccaca gccacaagaa gaacctgttc gagaccccca tcctggcccg gtacgtgcgg 2880

atcctgcccg tggcctggca caaccggatc gccctgcggc tggagctgct gggctgc 2937

<210> 113

<211> 2952

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 113

gacatctgcg accccaaccc ctgcgagaac ggcggcatct gcctgcccgg cctggccgac 60

ggcagcttca gctgcgagtg ccccgacggc ttcaccgacc ccaactgcag cagcgtggtg 120

gaggtggcca gcgacgagga ggagcccacc aacatcaacg agtgcgaggt ggagccctgc 180

aagaacggcg gcatctgcac cgacctggtg gccaactaca gctgcgagtg ccccggcgag 240

ttcatgggcc ggaactgcca gtacaaggac gcccacaaga gcgaggtggc ccaccggttc 300

aaggacctgg gcgaggagaa cttcaaggcc ctggtgctga tcgccttcgc ccagtacctg 350

cagcagagcc ccttcgagga ccacgtgaag ctggtgaacg aggtgaccga gttcgccaag 420

acctgcgtgg ccgacgagag cgccgagaac tgcgacaaga gcctgcacac cctgttcggc 480

gacaagctgt gcaccgtggc caccctgcgg gagacctacg gcgagatggc cgactgctgc 540

gccaagcagg agcccgagcg gaacgagtgc ttcctgcagc acaaggacga caaccccaac 600

ctgccccggc tggtgcggcc cgaggtggac gtgatgtgca ccgccttcca cgacaacgag 660

gagaccttcc tgaagaagta cctgtacgag atcgcccggc ggcaccccta cttctacgcc 720

cccgagctgc tgttcttcgc caagcggtac aaggccgcct tcaccgagtg ctgccaggcc 780

gccgacaagg ccgcctgcct gctgcccaag ctggacgagc tgcgggacga gggcaaggcc 840

agcagcgcca agcagcggct gaagtgcgcc agcctgcaga agttcggcga gcgggccttc 900

aaggcctggg ccgtggcccg gctgagccag cggttcccca aggccgagtt cgccgaggtg 960

agcaagctgg tgaccgacct gaccaaggtg cacaccgagt gctgccacgg cgacctgctg 1020

gagtgcgccg acgaccgggc cgacctggcc aagtacatct gcgagaacca ggacagcatc 1080

agcagcaagc tgaaggagtg ctgcgagaag cccctgctgg agaagagcca ctgcatcgcc 1140

gaggtggaga acgacgagat gcccgccgac ctgcccagcc tggccgccga cttcgtggag 1200

agcaaggacg tgtgcaagaa ctacgccgag gccaaggacg tgttcctggg catgttcctg 1260

tacgagtacg cccggcggca ccccgactac agcgtggtgc tgctgctgcg gctggccaag 1320

acctacgaga ccaccctgga gaagtgctgc gccgccgccg acccccacga gtgctacgcc 1380

aaggtgttcg acgagttcaa gcccctggtg gaggagcccc agaacctgat caagcagaac 1440

tgcgagctgt tcgagcagct gggcgagtac aagttccaga acgccctgct ggtgcggtac 1500

accaagaagg tgccccaggt gagcaccccc accctggtgg aggtgagccg gaacctgggc 1560

aaggtgggca gcaagtgctg caagcacccc gaggccaagc ggatgccctg cgccgaggac 1620

tacctgagcg tggtgctgaa ccagctgtgc gtgctgcacg agaagacccc cgtgagcgac 1680

cgggtgacca agtgctgcac cgagagcctg gtgaaccggc ggccctgctt cagcgccctg 1740

gaggtggacg agacctacgt gcccaaggag ttcaacgccg agaccttcac cttccacgcc 1800

gacatctgca ccctgagcga gaaggagcgg cagatcaaga agcagaccgc cctggtggag 1860

ctggtgaagc acaagcccaa ggccaccaag gagcagctga aggccgtgat ggacgacttc 1920

gccgccttcg tggagaagtg ctgcaaggcc gacgacaagg agacctgctt cgccgaggag 1980

ggcaagaagc tggtggcctg cgtggagccc ctgggcatgg agaacggcaa catcgccaac 2040

agccagatcg ccgccagcag cgtgcgggtg accttcctgg gcctgcagca ctgggtgccc 2100

gagctggccc ggctgaaccg ggccggcatg gtgaacgcct ggacccccag cagcaacgac 2160

gacaacccct ggatccaggt gaacctgctg cggcggatgt gggtgaccgg cgtggtgacc 2220

cagggcgcca gccggctggc cagccacgag tacctgaagg ccttcaaggt ggcctacagc 2280

ctgaacggcc acgagttcga cttcatccac gacgtgaaca agaagcacaa ggagttcgtg 2340

ggcaactgga acaagaacgc cgtgcacgtg aacctgttcg agacccccgt ggaggcccag 2400

tacgtgcggc tgtaccccac cagctgccac accgcctgca ccctgcggtt cgagctgctg 2460

ggctgcgagc tgaacggctg cgccaacccc ctgggcctga agaacaacag catccccgac 2520

aagcagatca ccgccagcag cagctacaag acctggggcc tgcacctgtt cagctggaac 2580

cccagctacg cccggctgga caagcagggc aacttcaacg cctgggtggc cggcagctac 2640

ggcaacgacc agtggctgca ggtggacctg ggcagcagca aggaggtgac cggcatcatc 2700

acccagggcg cccggaactt cggcagcgtg cagttcgtgg ccagctacaa ggtggcctac 2760

agcaacgaca gcgccaactg gaccgagtac caggaccccc ggaccggcag cagcaagatc 2820

ttccccggca actgggacaa ccacagccac aagaagaacc tgttcgagac ccccatcctg 2880

gcccggtacg tgcggatcct gcccgtggcc tggcacaacc ggatcgccct gcggctggag 2940

ctgctgggct gc 2952

<210> 114

<211> 5

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Peptit tổng hợp"

<400> 114

Gly Gly Gly Gly Ser

1 5

<210> 115

<211> 212

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 115

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu

1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys

20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Cys Val

35 40 45

Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile Ala

50 55 60

Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val Pro

65 70 75 80

Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr Pro

85 90 95

Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg Arg

100 105 110

Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala Ser

115 120 125

His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly His

130 135 140

Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe Val

145 150 155 160

Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr Pro

165 170 175

Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr Ala

180 185 190

Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly His His

195 200 205

His His His His

210

<210> 116

<211> 636

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 116

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaatg tgtggaaccc ctcggcatgg aaaacggcaa tatcgccaat 180

agccagatcg ccgccagcag cgtcagagtg acatttctgg gactgcaaca ctgggtgcca 240

gagctggcca gactgaatag agccggcatg gttaacgcct ggacacccag cagcaacgac 300

gacaacccct ggattcaagt gaacctgctg cggcgtatgt gggtcacagg tgttgttaca 360

cagggcgcaa gcagactggc cagccacgag tatctgaagg cctttaaggt ggcctacagc 420

ctgaacggcc acgagttcga cttcatccac gacgtgaaca agaagcacaa agagttcgtc 480

ggcaactgga acaagaacgc cgtgcacgtg aacctgttcg agacacctgt ggaagcccag 540

tacgtgcggc tgtaccctac aagctgtcac accgcctgca cactgagatt cgagctgctg 600

ggctgtgaac tgaatggcca ccaccaccat caccac 636

<210> 117

<211> 805

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 117

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80
Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95
Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
100 105 110
Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
115 120 125
Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
130 135 140
Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
145 150 155 160
Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
165 170 175
Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
180 185 190
Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
195 200 205
Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
210 215 220
Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
225 230 235 240
Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
245 250 255
Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
260 265 270
Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
275 280 285
Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
290 295 300
Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser

305 310 315 320

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His

325 330 335

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

340 345 350

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

355 360 365

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg

370 375 380

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr

385 390 395 400

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu

405 410 415

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

420 425 430

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu

435 440 445

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro

450 455 460

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys

465 470 475 480

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys

485 490 495

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His

500 505 510

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser

515 520 525

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr

530 535 540

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp

545 550 555 560

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
565 570 575
Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
580 585 590
Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
595 600 605
Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
610 615 620
Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly
625 630 635 640
Gly Ser Gly Gly Ser Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn
645 650 655
Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu
660 665 670
Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly
675 680 685
Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile
690 695 700
Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr Gly Val Val Thr Gln
705 710 715 720
Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val
725 730 735
Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn
740 745 750
Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His
755 760 765
Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr
770 775 780
Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly
785 790 795 800
Cys Glu Leu Asn Gly

805

<210> 118

<211> 2415

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 118

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atctgatgcc cacaagagcg aggtggccca cagattcaag 180

gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240

cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300

tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360

aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420

aagcaagagc ccgagagaaa cgagtgcttc ctccagcaca aggacgacaa ccccaacctg 480

cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540

accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgcccct 600

gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa atgcgccagc ctccagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgcaccc tgtccgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaggctgct ctcggacttg gtggaagcgg aggaagtggt 1920

ggatctggcg gatcttgtgt ggaacccctc ggcatggaaa acggcaatat cgccaatagc 1980

cagattgccg ccagcagcgt cagagtgaca tttctgggac tgcaacactg ggtgcccgag 2040

ctggctagac tgaatagagc cggcatggtc aacgcctgga cacccagcag caacgacgat 2100

aatccctgga ttcaagtgaa cctgctgcgg cgtatgtggg tcacaggtgt tgttacacag 2160

ggcgcaagca gactggccag ccacgagtat ctgaaggcct ttaaggtggc ctacagcctg 2220

aacggccacg agttcgactt catccacgac gtgaacaaga agcacaaaga gtttgtcggc 2280

aactggaaca agaacgccgt gcacgtgaac ctgttcgaga cacctgtgga agcccagtac 2340

gtgcggctgt accctacaag ctgtcacacc gcctgcactc tgagattcga actgctggga 2400

tgcgagctga acggc 2415

<210> 119

<211> 789

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 119

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu

1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys

20 25 30

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

275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala

290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys

305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile

325 330 335

Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala

340 345 350

Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala

355 360 365

Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His

370 375 380

Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu

385 390 395 400

Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr

405 410 415

Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn

420 425 430

Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys

435 440 445

Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val

450 455 460

Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly

465 470 475 480

Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu

485 490 495

Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys

500 505 510

Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val

515 520 525

Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val
530 535 540

Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys
545 550 555 560

Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val
565 570 575

Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala
580 585 590

Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp
595 600 605

Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala
610 615 620

Ser Gln Ala Ala Leu Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn
625 630 635 640

Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu
645 650 655

Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly
660 665 670

Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile
675 680 685

Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr Gly Val Val Thr Gln
690 695 700

Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val
705 710 715 720

Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn
725 730 735

Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His
740 745 750

Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr
755 760 765

Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly

770 775 780

Cys Glu Leu Asn Gly

785

<210> 120

<211> 2367

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 120

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaagga tgcccacaag agcgaggtgg cccacagatt caaggatctg 180

ggcgaagaga acttcaaggc cctggtgctg atcgccttcg ctcagtatct ccagcagagc 240

cctttcgagg accacgtgaa gctggtcaac gaagtgaccg agttcgccaa gacctgtgtg 300

gccgatgaga gcgccgagaa ctgtgataag agcctgcaca ccctgttcgg cgacaagctg 360

tgtacagtgg ccacactgag agaaacctac ggcgagatgg ccgactgctg tgccaagcaa 420

gagcccgaga gaaacgagtg cttcctccag cacaaggacg acaaccccaa cctgcctaga 480

ctcgtgcgac ccgaagtgga tgtgatgtgc accgcctttc acgacaacga ggaaaccttc 540

ctgaagaagt acctgtacga gatcgccaga cggcacccct acttttatgc ccctgagctg 600

ctgttcttcg ccaagcggta taaggccgcc ttcaccgaat gttgccaggc cgctgataag 660

gctgcctgtc tgctgcctaa gctggacgag ctgagagatg agggcaaagc cagctctgcc 720

aagcagagac tgaaatgcgc cagcctccag aagttcggcg agagagcttt taaggcctgg 780

gccgttgcca gactgagcca gagatttcct aaggccgagt ttgccgaggt gtccaagctc 840

gtgaccgatc tgacaaaggt gcacaccgag tgctgtcacg gcgatctgct ggaatgtgcc 900

gacgatagag ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 960

ctgaaagagt gctgcgagaa gcccctgctg gaaaagtctc actgtatcgc cgaggtggaa 1020

aacgacgaga tgcctgccga tctgcctagc ctggctgccg atttcgtgga aagcaaggac 1080

gtgtgcaaga actacgccga ggccaaggat gtgtttctgg gcatgtttct gtatgagtac 1140

gcccgcagac accccgacta ttctgtggtt ctgctgctgc ggctggccaa aacctacgag 1200

acaaccctgg aaaaatgctg cgccgctgcc gatcctcacg agtgttatgc caaggtgttc 1260

gacgagttca agcctctggt ggaagaaccc cagaacctga tcaagcagaa ctgcgagctg 1320

ttcgagcagc tgggcgagta caagttccag aatgccctgc tcgtgcggta caccaagaaa 1380

gtgcctcagg tgtccacacc tacactggtt gaggtgtccc ggaatctggg caaagtgggc 1440

agcaagtgtt gcaagcaccc tgaggccaag agaatgcctt gcgccgagga ttacctgagc 1500

gtggtgctga atcagctgtg cgtgctgcac gagaaaaccc ctgtgtccga cagagtgacc 1560

aagtgctgta ccgagagcct cgtgaacaga aggccttgct ttagcgccct ggaagtggac 1620

gagacatacg tgcccaaaga gttcaacgcc gagacattca ccttccacgc cgacatctgc 1680

accctgtccg agaaagagcg gcagatcaag aagcagacag ccctggtcga gctggttaag 1740

cacaagccca aggccaccaa agaacagctg aaggccgtga tggacgactt cgccgccttt 1800

gtcgagaagt gctgcaaggc cgacgacaaa gagacatgct tcgccgaaga gggcaagaaa 1860

ctggtggctg cttctcaggc cgctctgtgt gtggaacccc tcggcatgga aaacggcaat 1920

atcgccaata gccagattgc cgccagcagc gtcagagtga catttctggg actgcaacac 1980

tgggtgcccg agctggctag actgaataga gccggcatgg tcaacgcctg gacacccagc 2040

agcaacgacg ataatccctg gattcaagtg aacctgctgc ggcgtatgtg ggtcacaggt 2100

gttgttacac agggcgcaag cagactggcc agccacgagt atctgaaggc ctttaaggtg 2160

gcctacagcc tgaacggcca cgagttcgac ttcatccacg acgtgaacaa gaagcacaaa 2220

gagtttgtcg gcaactggaa caagaacgcc gtgcacgtga acctgttcga gacacctgtg 2280

gaagcccagt acgtgcggct gtaccctaca agctgtcaca ccgcctgcac tctgagattc 2340

gaactgctgg gatgcgagct gaacggc 2367

<210> 121

<211> 783

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 121

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu

1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys

20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala

35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn

50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser

65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala

85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu

100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu

115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg

130 135 140

Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg

145 150 155 160

Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn

165 170 175

Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His

180 185 190

Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys

195 200 205

Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu

210 215 220

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

225 230 235 240

Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala

245 250 255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala

260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His
275 280 285
Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala
290 295 300
Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys
305 310 315 320
Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile
325 330 335
Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala
340 345 350
Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala
355 360 365
Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His
370 375 380
Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu
385 390 395 400
Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr
405 410 415
Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn
420 425 430
Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys
435 440 445
Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val
450 455 460
Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly
465 470 475 480
Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu
485 490 495
Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys
500 505 510
Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val

515 520 525

Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val

530 535 540

Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys

545 550 555 560

Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val

565 570 575

Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala

580 585 590

Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp

595 600 605

Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Cys

610 615 620

Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile

625 630 635 640

Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val

645 650 655

Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr

660 665 670

Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg

675 680 685

Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala

690 695 700

Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly

705 710 715 720

His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe

725 730 735

Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr

740 745 750

Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr

755 760 765

Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly

770 775 780

<210> 122

<211> 2349

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 122

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaagga tgcccacaag agcgaggtgg cccacagatt caaggatctg 180

ggcgaagaga acttcaaggc cctggtgctg atcgccttcg ctcagtatct ccagcagagc 240

cctttcgagg accacgtgaa gctggtcaac gaagtgaccg agttcgccaa gacctgtgtg 300

gccgatgaga gcgccgagaa ctgtgataag agcctgcaca ccctgttcgg cgacaagctg 360

tgtacagtgg ccacactgag agaaacctac ggcgagatgg ccgactgctg tgccaagcaa 420

gagcccgaga gaaacgagtg cttcctccag cacaaggacg acaaccccaa cctgcctaga 480

ctcgtgcgac ccgaagtgga tgtgatgtgc accgcctttc acgacaacga ggaaaccttc 540

ctgaagaagt acctgtacga gatcgccaga cggcaccccт actttatgc ccctgagctg 600

ctgttcttcg ccaagcggta taaggccgcc ttcaccgaat gttgccaggc cgctgataag 660

gctgcctgtc tgctgcctaa gctggacgag ctgagagatg agggcaaagc cagctctgcc 720

aagcagagac tgaaatgcgc cagcctccag aagttcggcg agagagcttt taaggcctgg 780

gccgttgcca gactgagcca gagatttcct aaggccgagt ttgccgaggt gtccaagctc 840

gtgaccgatc tgacaaaggt gcacaccgag tgctgtcacg gcgatctgct ggaatgtgcc 900

gacgatagag ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 960

ctgaaagagt gctgcgagaa gcccctgctg gaaaagtctc actgtatcgc cgaggtggaa 1020

aacgacgaga tgcctgccga tctgcctagc ctggctgccg atttcgtgga aagcaaggac 1080

gtgtgcaaga actacgccga ggccaaggat gtgtttctgg gcatgtttct gtatgagtac 1140

gcccgcagac accccgacta ttctgtggtt ctgctgctgc ggctggccaa aacctacgag 1200

acaaccctgg aaaaatgctg cgccgctgcc gatcctcacg agtgttatgc caaggtgttc 1260

gacgagttca agcctctggt ggaagaaccc cagaacctga tcaagcagaa ctgcgagctg 1320

ttcgagcagc tgggcgagta caagttccag aatgccctgc tcgtgcggta caccaagaaa 1380

gtgcctcagg tgtccacacc tacactggtt gaggtgtccc ggaatctggg caaagtgggc 1440

agcaagtgtt gcaagcaccc tgaggccaag agaatgcctt gcgccgagga ttacctgagc 1500

gtggtgctga atcagctgtg cgtgctgcac gagaaaaccc ctgtgtccga cagagtgacc 1560

aagtgctgta ccgagagcct cgtgaacaga aggccttgct ttagcgccct ggaagtggac 1620

gagacatacg tgcccaaaga gttcaacgcc gagacattca ccttccacgc cgacatctgc 1680

accctgtccg agaaagagcg gcagatcaag aagcagacag ccctggtcga gctggttaag 1740

cacaagccca aggccaccaa agaacagctg aaggccgtga tggacgactt cgccgccttt 1800

gtcgagaagt gctgcaaggc cgacgacaaa gagacatgct tcgccgaaga gggcaagaaa 1860

ctggtggcct gtgtggaacc cctcggcatg gaaaacggca atatcgccaa tagccagatt 1920

gccgccagca gcgtcagagt gacatttctg ggactgcaac actgggtgcc cgagctggct 1980

agactgaata gagccggcat ggtcaacgcc tggacaccca gcagcaacga cgataatccc 2040

tggattcaag tgaacctgct gcggcgtatg tgggtcacag gtgttgttac acagggcgca 2100

agcagactgg ccagccacga gtatctgaag gcctttaagg tggcctacag cctgaacggc 2160

cacgagttcg acttcatcca cgacgtgaac aagaagcaca aagagtttgt cggcaactgg 2220

aacaagaacg ccgtgcacgt gaacctgttc gagacacctg tggaagccca gtacgtgcgg 2280

ctgtacccta caagctgtca caccgcctgc actctgagat tcgaactgct gggatgcgag 2340

ctgaacggc 2349

<210> 123

<211> 227

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 123

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly

1 5 10 15

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met

20 25 30

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

<210> 124
<211> 681
<212> ADN
<213> Trình tự Nhân tạo
<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 124

gataagaccc acacctgtcc tccatgtcct gctccagaac tgctcggcgg accctccgtt 60

ttcctgtttc cacctaagcc taaggacacc ctgatgatca gcagaacccc tgaagtgacc 120

tgtgtggtgg tggccgtgtc tcacgaagat cccgaagtga agttcaattg gtacgtggac 180

ggcgtggaag tgcacaacgc caagaccaag cctagagagg aacagtacaa cagcacctac 240

agagtggtgt ccgtgctgac cgtgctgcac caggattggc tgaacggcaa agagtacaag 300

tgcaaggtgt ccaacaaggc cctggccgct cctatcgaga aaaccatctc taaggccaag 360

ggccagcctc gggaacctca agtctgtaca ctgcctccta gccgggacga gctgaccaaa 420

aatcaggtgt ccctgagctg cgccgtgaag ggcttttacc cttccgatat cgccgtggaa 480

tgggagagca atggccagcc tgagaacaac tacaagacca cacctcctgt gctggacagc 540

gacggctcat tctttctggt gtccaagctg acagtggaca agagcagatg gcagcagggc 600

aacgtgttca gctgttctgt gatgcacgag gccctgcaca accactacac ccagaagtct 660

ctgtctctga gccccggcaa a 681

<210> 125

<211> 605

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 125

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
50 55 60

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
65 70 75 80
Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Ala Val Ser His
85 90 95
Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
100 105 110
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
115 120 125
Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
130 135 140
Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Ala Ala Pro Ile
145 150 155 160
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
165 170 175
Tyr Thr Leu Pro Pro Cys Arg Glu Glu Met Thr Lys Asn Gln Val Ser
180 185 190
Leu Trp Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
195 200 205
Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
210 215 220
Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
225 230 235 240
Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
245 250 255
His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
260 265 270
Pro Gly Lys Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Cys
275 280 285
Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile
290 295 300
Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val

305 310 315 320

Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr

325 330 335

Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg

340 345 350

Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala

355 360 365

Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly

370 375 380

His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe

385 390 395 400

Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr

405 410 415

Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr

420 425 430

Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys

435 440 445

Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile

450 455 460

Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp

465 470 475 480

Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp

485 490 495

Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly

500 505 510

Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe

515 520 525

Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp

530 535 540

Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys

545 550 555 560

Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe
565 570 575
Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp
580 585 590
His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys
595 600 605

<210> 126

<211> 1815

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 126

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaggg cagcgataag acccacacct gtcctccatg tcctgctcca 180

gaactgctcg gcggaccctc cgttttcctg tttccaccta agcctaagga caccctgatg 240

atcagcagaa cccctgaagt gacctgtgtg gtggtggccg tgtctcacga agatcccgaa 300

gtgaagttca attggtacgt ggacggcgtg gaagtgcaca acgccaagac caagcctaga 360

gaggaacagt acaacagcac ctacagagtg gtgtccgtgc tgaccgtgct gcaccaggat 420

tggctgaacg gcaaagagta caagtgcaag gtgtccaaca aggccctggc cgctcctatc 480

gagaaaacca tctctaaggc caagggccag cctcgggaac ctcaggttta caccctgcct 540

ccatgccggg aagagatgac caagaatcag gtgtccctgt ggtgcctggt caagggcttc 600

taccttccg atatcgccgt ggaatgggag agcaatggcc agcctgagaa caactacaag 660

accacacctc ctgtgctgga cagcgacggc tcattcttcc tgtacagcaa gctgacagtg 720

gacaagagca gatggcagca gggcaacgtg ttcagctgtt ctgtgatgca cgaggccctg 780

cacaaccact acacccagaa gtctctgtct ctgagccctg gcaaaggcgg aagcggtgga 840

agcggaggat ctggcggatc ttgtgtggaa cccctcggca tggaaaacgg caatatcgcc 900

aatagccaga tcgccgccag cagcgtcaga gtgacatttc tgggactgca acactgggtg 960

ccagagctgg ccagactgaa tagagccggc atggttaacg cctggacacc cagcagcaac 1020

gacgacaacc cctggattca agtgaacctg ctgcggcgta tgtgggtcac aggtgttgtt 1080

acacagggcg caagcagact ggccagccac gagtatctga aggcctttaa ggtggcctac 1140

agcctgaacg gccacgagtt cgacttcatc cacgacgtga acaagaagca caaagagttc 1200

gtcggcaact ggaacaagaa cgccgtgcac gtgaacctgt tcgagacacc tgtggaagcc 1260

cagtacgtgc ggctgtaccc tacaagctgt cacaccgcct gcacactgag attcgagctg 1320

ctgggctgcg agctgaatgg ctgtgctaat cctctgggcc tgaagaacaa tagcatcccc 1380

gacaagcaga tcaccgcctc cagcagctat aagacatggg gcctgcacct gtttagctgg 1440

aaccctagct acgccagact ggacaagcag ggaaacttca atgcctgggt ggccggcagc 1500

tacggcaatg atcaatggct gcaagtggac ctgggcagca gcaaagaagt gaccggcatc 1560

attacccagg gcgctagaaa tttcggcagc gtgcagttcg tggccagcta caaagtggcc 1620

tactccaacg acagcgccaa ctggaccgag tatcaggacc ctagaaccgg cagctccaag 1680

atcttccccg gcaattggga caaccacagc cacaagaaga atctgttcga aacccctatc 1740

ctggccagat atgtgcgcat tctgcccgtg gcctggcaca acagaattgc cctgagactg 1800

gaactgctgg gatgc 1815

<210> 127

<211> 227

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 127

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly

1 5 10 15

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met

20 25 30

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Ala Val Ser His

35 40 45

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val

50 55 60

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr

65 70 75 80

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly

85 90 95

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Ala Ala Pro Ile

100 105 110

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

115 120 125

Tyr Thr Leu Pro Pro Cys Arg Glu Glu Met Thr Lys Asn Gln Val Ser

130 135 140

Leu Trp Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu

145 150 155 160

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro

165 170 175

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val

180 185 190

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met

195 200 205

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser

210 215 220

Pro Gly Lys

225

<210> 128

<211> 681

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 128

gataagaccc acacctgtcc tccatgtcct gctccagaac tgctcggcgg accctccgtt 60

ttcctgtttc cacctaagcc taaggacacc ctgatgatca gcagaacccc tgaagtgacc 120

tgtgtggtgg tggccgtgtc tcacgaagat cccgaagtga agttcaattg gtacgtggac 180

ggcgtggaag tgcacaacgc caagaccaag cctagagagg aacagtacaa cagcacctac 240

agagtggtgt ccgtgctgac cgtgctgcac caggattggc tgaacggcaa agagtacaag 300

tgcaaggtgt ccaacaaggc cctggccgct cctatcgaga aaaccatctc taaggccaag 360

ggccagcctc gggaacctca ggtttacacc ctgcctccat gccgggaaga gatgaccaag 420

aatcaggtgt ccctgtggtg cctggtcaag ggcttctacc cttccgatat cgccgtggaa 480

tgggagagca atggccagcc tgagaacaac tacaagacca cacctcctgt gctggacagc 540

gacggctcat tcttcctgta cagcaagctg acagtggaca agagcagatg gcagcagggc 600

aacgtgttca gctgttctgt gatgcacgag gccctgcaca accactacac ccagaagtct 660

ctgtctctga gccccggcaa a 681

<210> 129

<211> 605

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 129

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
50 55 60

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
65 70 75 80

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Ala Val Ser His
85 90 95

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val

100 105 110

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr

115 120 125

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly

130 135 140

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Ala Ala Pro Ile

145 150 155 160

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

165 170 175

Cys Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser

180 185 190

Leu Ser Cys Ala Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu

195 200 205

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro

210 215 220

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Val Ser Lys Leu Thr Val

225 230 235 240

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met

245 250 255

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser

260 265 270

Pro Gly Lys Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Cys

275 280 285

Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln Ile

290 295 300

Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val

305 310 315 320

Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr

325 330 335

Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg

340 345 350

Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala
355 360 365

Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly
370 375 380

His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe
385 390 395 400

Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr
405 410 415

Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr
420 425 430

Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly Cys
435 440 445

Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile
450 455 460

Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp
465 470 475 480

Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp
485 490 495

Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly
500 505 510

Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe
515 520 525

Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp
530 535 540

Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys
545 550 555 560

Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe
565 570 575

Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp
580 585 590

His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys

595 600 605

<210> 130

<211> 1815

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 130

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaggg cagcgataag acccacacct gtcctccatg tcctgctcca 180

gaactgctcg gcggaccctc cgttttcctg tttccaccta agcctaagga caccctgatg 240

atcagcagaa cccctgaagt gacctgtgtg gtggtggccg tgtctcacga agatcccgaa 300

gtgaagttca attggtacgt ggacggcgtg gaagtgcaca acgccaagac caagcctaga 360

gaggaacagt acaacagcac ctacagagtg gtgtccgtgc tgaccgtgct gcaccaggat 420

tggctgaacg gcaaagagta caagtgcaag gtgtccaaca aggccctggc cgctcctatc 480

gagaaaacca tctctaaggc caagggccag cctcgggaac ctcaagtctg tacactgcct 540

cctagccggg acgagctgac caaaaatcag gtgtccctga gctgcgccgt gaagggcttt 600

taccettccg atatcgccgt ggaatgggag agcaatggcc agcctgagaa caactacaag 660

accacacctc ctgtgctgga cagcgacggc tcattctttc tggtgtccaa gctgacagtg 720

gacaagagca gatggcagca gggcaacgtg ttcagctgtt ctgtgatgca cgaggccctg 780

cacaaccact acacccagaa gtctctgtct ctgagccctg gcaaaggcgg aagcggtgga 840

agcggaggat ctggcggatc ttgtgtggaa cccctcggca tggaaaacgg caatatcgcc 900

aatagccaga tcgccgccag cagcgtcaga gtgacatttc tgggactgca acactgggtg 960

ccagagctgg ccagactgaa tagagccggc atggttaacg cctggacacc cagcagcaac 1020

gacgacaacc cctggattca agtgaacctg ctgcggcgta tgtgggtcac aggtgttgtt 1080

acacagggcg caagcagact ggccagccac gagtatctga aggcctttaa ggtggcctac 1140

agcctgaacg gccacgagtt cgacttcatc cacgacgtga acaagaagca caaagagttc 1200

gtcggcaact ggaacaagaa cgccgtgcac gtgaacctgt tcgagacacc tgtggaagcc 1260

cagtacgtgc ggctgtaccc tacaagctgt cacaccgcct gcacactgag attcgagctg 1320

ctgggctgcg agctgaatgg ctgtgctaat cctctgggcc tgaagaacaa tagcatcccc 1380

gacaagcaga tcaccgcctc cagcagctat aagacatggg gcctgcacct gtttagctgg 1440

aaccctagct acgccagact ggacaagcag ggaaacttca atgcctgggt ggccggcagc 1500

tacggcaatg atcaatggct gcaagtggac ctgggcagca gcaaagaagt gaccggcatc 1560

attacccagg gcgctagaaa tttcggcagc gtgcagttcg tggccagcta caaagtggcc 1620

tactccaacg acagcgccaa ctggaccgag tatcaggacc ctagaaccgg cagctccaag 1680

atcttccccg gcaattggga caaccacagc cacaagaaga atctgttcga aacccctatc 1740

ctggccagat atgtgcgcat tctgcccgtg gcctggcaca acagaattgc cctgagactg 1800

gaactgctgg gatgc 1815

<210> 131

<211> 783

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 131

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Glu Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala
35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn
50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser
65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala
85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu
100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu
115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg
130 135 140

Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg
145 150 155 160

Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn
165 170 175

Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His
180 185 190

Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys
195 200 205

Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu
210 215 220

Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala
225 230 235 240

Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala
245 250 255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala
260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His
275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala
290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys
305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile
325 330 335

Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala
340 345 350

Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala

355 360 365

Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His

370 375 380

Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu

385 390 395 400

Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr

405 410 415

Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn

420 425 430

Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys

435 440 445

Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val

450 455 460

Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly

465 470 475 480

Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu

485 490 495

Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys

500 505 510

Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val

515 520 525

Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val

530 535 540

Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys

545 550 555 560

Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val

565 570 575

Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala

580 585 590

Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp

595 600 605

Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Cys
610 615 620

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

Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp Val
645 650 655

Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp Thr
660 665 670

Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu Arg
675 680 685

Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu Ala
690 695 700

Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn Gly
705 710 715 720

His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu Phe
725 730 735

Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu Thr
740 745 750

Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His Thr
755 760 765

Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly
770 775 780

<210> 132

<211> 2349

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 132

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatcagtcaa 60

gaagtgcggg gcgaagtctt tcccagctac acctgtacct gtctgaaggg ctatgccggc 120

aaccactgcg agacaaagga tgcccacaag agcgaggtgg cccacagatt caaggatctg 180

ggcgaagaga acttcaaggc cctggtgctg atcgccttcg ctcagtatct ccagcagagc 240

cctttcgagg accacgtgaa gctggtcaac gaagtgaccg agttcgccaa gacctgtgtg 300

gccgatgaga gcgccgagaa ctgtgataag agcctgcaca ccctgttcgg cgacaagctg 360

tgtacagtgg ccacactgag agaaacctac ggcgagatgg ccgactgctg tgccaagcaa 420

gagcccgaga gaaacgagtg cttcctccag cacaaggacg acaaccccaa cctgcctaga 480

ctcgtgcgac ccgaagtgga tgtgatgtgc accgcctttc acgacaacga ggaaaccttc 540

ctgaagaagt acctgtacga gatcgccaga cggcaccct actttatgc ccctgagctg 600

ctgttcttcg ccaagcggta taaggccgcc ttcaccgaat gttgccaggc cgctgataag 660

gctgcctgtc tgctgcctaa gctggacgag ctgagagatg agggcaaagc cagctctgcc 720

aagcagagac tgaaatgcgc cagcctccag aagttcggcg agagagcttt taaggcctgg 780

gccgttgcca gactgagcca gagatttcct aaggccgagt ttgccgaggt gtccaagctc 840

gtgaccgatc tgacaaaggt gcacaccgag tgctgtcacg gcgatctgct ggaatgtgcc 900

gacgatagag ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 960

ctgaaagagt gctgcgagaa gcccctgctg gaaaagtctc actgtatcgc cgaggtggaa 1020

aacgacgaga tgcctgccga tctgcctagc ctggctgccg atttcgtgga aagcaaggac 1080

gtgtgcaaga actacgccga ggccaaggat gtgtttctgg gcatgtttct gtatgagtac 1140

gcccgcagac accccgacta ttctgtggtt ctgctgctgc ggctggccaa aacctacgag 1200

acaaccctgg aaaaatgctg cgccgctgcc gatcctcacg agtgttatgc caaggtgttc 1260

gacgagttca agcctctggt ggaagaaccc cagaacctga tcaagcagaa ctgcgagctg 1320

ttcgagcagc tgggcgagta caagttccag aatgccctgc tcgtgcggta caccaagaaa 1380

gtgcctcagg tgtccacacc tacactggtt gaggtgtccc ggaatctggg caaagtgggc 1440

agcaagtgtt gcaagcaccc tgaggccaag agaatgcctt gcgccgagga ttatctgagc 1500

gtggtgctga atcagctgtg cgtgctgcac gagaaaaccc ctgtgtccga cagagtgacc 1560

aagtgctgta ccgagagcct cgtgaacaga aggccttgct ttagcgccct ggaagtggac 1620

gagacatacg tgcccaaaga gttcaacgcc gagacattca ccttccacgc cgacatctgc 1680

accctgtccg agaaagagcg gcagatcaag aagcagacag ccctggtcga gctggttaag 1740

cacaagccca aggccaccaa agaacagctg aaggccgtga tggacgactt cgccgccttt 1800

gtcgagaagt gctgcaaggc cgacgacaaa gagacatgct tcgccgaaga gggcaagaaa 1860

ctggtggcct gtgtggaacc cctcggcatg gaaaacggca atatcgccaa tagccagatt 1920

gccgccagca gcgtcagagt gacatttctg ggactgcaac actgggtgcc cgagctggct 1980

agactgaata gagccggcat ggtcaacgcc tggacaccca gcagcaacga cgataatccc 2040

tggattcaag tgaacctgct gcggcgtatg tgggtcacag gtgttgttac acagggcgca 2100

agcagactgg ccagccacga gtatctgaag gcctttaagg tggcctacag cctgaacggc 2160

cacgagttcg acttcatcca cgacgtgaac aagaagcaca aagagtttgt cggcaactgg 2220

aacaagaacg ccgtgcacgt gaacctgttc gagacacctg tggaagccca gtacgtgcgg 2280

ctgtacccta caagctgtca caccgcctgc actctgagat tcgaactgct gggatgcgag 2340

ctgaacggc 2349

<210> 133

<211> 872

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 133

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro

1 5 10 15

Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr

20 25 30

Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu

35 40 45

Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly

50 55 60

Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr

65 70 75 80

Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn

85 90 95

Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr

100 105 110

Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly

115 120 125

Arg Asn Cys Gln Tyr Lys Asp Ala His Lys Ser Glu Val Ala His Arg

130 135 140

Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala

145 150 155 160

Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu

165 170 175

Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser

180 185 190

Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu

195 200 205

Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys

210 215 220

Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys

225 230 235 240

Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val

245 250 255

Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr

260 265 270

Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu

275 280 285

Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln

290 295 300

Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg

305 310 315 320

Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser

325 330 335

Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg

340 345 350

Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu

355 360 365

Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu
370 375 380
Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu
385 390 395 400
Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro
405 410 415
Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met
420 425 430
Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp
435 440 445
Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe
450 455 460
Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu
465 470 475 480
Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala
485 490 495
Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys
500 505 510
Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu
515 520 525
Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg
530 535 540
Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val
545 550 555 560
Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu
565 570 575
Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn
580 585 590
Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr
595 600 605
Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala

610 615 620

Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr

625 630 635 640

Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln

645 650 655

Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys

660 665 670

Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe

675 680 685

Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu

690 695 700

Glu Gly Lys Lys Leu Val Ala Cys Ser Gly Pro Leu Gly Ile Glu Gly

705 710 715 720

Gly Ile Ile Ser Asn Gln Gln Ile Thr Ala Ser Ser Thr His Arg Ala

725 730 735

Leu Phe Gly Leu Gln Lys Trp Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys

740 745 750

Lys Gly Leu Ile Asn Ala Trp Thr Ala Ala Glu Asn Asp Arg Trp Pro

755 760 765

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

770 775 780

Thr Gln Gly Ala Lys Arg Ile Gly Ser Pro Glu Tyr Ile Lys Ser Tyr

785 790 795 800

Lys Ile Ala Tyr Ser Asn Asp Gly Lys Thr Trp Ala Met Tyr Lys Val

805 810 815

Lys Gly Thr Asn Glu Asp Met Val Phe Arg Gly Asn Ile Asp Asn Asn

820 825 830

Thr Pro Tyr Ala Asn Ser Phe Thr Pro Pro Ile Lys Ala Gln Tyr Val

835 840 845

Arg Leu Tyr Pro Gln Val Cys Arg Arg His Cys Thr Leu Arg Met Glu

850 855 860

Leu Leu Gly Cys Glu Leu Ser Gly

865 870

<210> 134

<211> 2616

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 134

gacatctgcg accccaatcc ttgcgagaat ggcggcattt gtctgcctgg actggccgat 60

ggcagcttct cttgtgaatg ccccgatggc ttcacagacc ccaattgcag ctctgtggtg 120

gaagtggcca gcgacgagga agaacctaca agcgctggcc cctgcacacc caatccatgt 180

cataatggcg gaacctgcga gatcagcgag gcctacagag gcgatacctt catcggctac 240

gtgtgcaagt gccccagagg cttcaatggc atccactgcc agcacaacat caacgagtgc 300

gaggtggaac catgcaagaa cggcggcatc tgtaccgacc tggtggccaa ttactcttgc 360

gagtgccctg gcgagttcat gggcagaaac tgccagtaca aggacgccca caagagcgag 420

gtggcccaca gattcaagga cctgggcgaa gagaacttca aggccctggt gctgatcgcc 480

ttcgctcagt atctccagca gagcccttтc gaggaccacg tgaagctggt caacgaagtg 540

accgagttcg ccaagacctg tgtggccgat gagagcgccg agaactgtga caagagcctg 600

cacacactgt tcggcgacaa gctgtgtacc gtggccacac tgagagaaac ctacggcgag 660

atggccgact gctgtgccaa gcaagagccc gagagaaacg agtgcttcct ccagcacaag 720

gatgacaacc ccaacctgcc tagactcgtg cggcctgaag tggatgtgat gtgcaccgcc 780

tttcacgaca acgaggaaac cttcctgaag aagtacctgt acgagatcgc cagacggcac 840

ccctactttt atgcccctga gctgctgttc ttcgccaagc ggtataaggc cgccttcacc 900

gaatgttgcc aggccgctga taaggctgcc tgtctgctgc ctaagctgga cgagctgaga 960

gatgagggca aagccagctc tgccaagcag agactgaaat gcgccagcct ccagaagttc 1020

ggcgagagag cttttaaggc ctgggccgtt gccagactga gccagagatt tcctaaggcc 1080

gagtttgccg aggtgtccaa gctcgtgacc gatctgacaa aggtgcacac cgagtgctgt 1140

cacggcgatc tgctggaatg tgccgacgat agagccgacc tggccaagta tatctgcgag 1200

aaccaggaca gcatcagcag caagctgaaa gagtgctgcg agaagcccct gctggaaaag 1260

tctcactgta tcgccgaagt ggaaaacgac gagatgcccg ccgatctgcc ttctctggct 1320

gccgatttcg tggaaagcaa ggatgtgtgc aagaactacg ccgaggccaa agatgtgttt 1380

ctgggcatgt ttctgtatga gtacgcccgc agacaccccg actattctgt ggttctgctg 1440

ctgcggctgg ccaagacata cgagacaacc ctggaaaaat gctgcgccgc tgccgatcct 1500

cacgagtgtt atgccaaggt gttcgacgag ttcaagccac tggtggaaga accccagaac 1560

ctgatcaagc agaactgcga gctgttcgag cagctgggcg agtacaagtt ccagaatgcc 1620

ctgctcgtgc ggtacaccaa gaaagtgcct caggtgtcca cacctacact ggttgaggtg 1680

tcccggaatc tgggcaaagt gggcagcaag tgttgcaagc accctgaggc caagagaatg 1740

ccttgcgccg aggattacct gagcgtggtg ctgaatcagc tgtgcgtgct gcacgagaaa 1800

acccctgtgt ccgacagagt gaccaagtgc tgtaccgaga gcctcgtgaa cagaaggcct 1860

tgctttagcg ccctggaagt ggacgagaca tacgtgccca aagagttcaa cgccgagaca 1920

ttcaccttcc acgccgatat ctgcaccctg tccgagaaag agcggcagat caagaagcag 1980

acagccctgg tcgagctggt taagcacaag cccaaggcca ccaaagaaca gctgaaggcc 2040

gtgatggacg acttcgccgc ctttgtcgag aagtgctgca aggccgacga caaagagaca 2100

tgcttcgccg aagagggcaa gaaactggtg gcctgttctg gccctctggg catcgaaggc 2160

ggcatcatca gcaatcagca gatcaccgcc agcagcaccc acagagcact gtttggcctg 2220

caaaagtggt atccctacta cgcccggctg aacaagaagg gcctgattaa cgcctggaca 2280

gccgccgaga atgacagatg gccctggatt cagatcaacc tccagcggaa gatgagagtg 2340

accggcgtta tcacacaggg cgcaaagaga atcggctccc ctgagtacat caagagctac 2400

aagatcgcct acagcaacga cggcaagacc tgggccatgt acaaagtgaa gggcaccaac 2460

gaggacatgg tgttccgggg caacatcgac aacaacaccc cttacgccaa cagcttcacc 2520

cctcctatca aggcccagta cgtgcggctg taccctcaag tgtgcagaag gcactgtacc 2580

ctgagaatgg aactgctggg ctgcgaactg tctggc 2616

<210> 135

<211> 869

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 135

Asp Ile Cys Asp Pro Asn Pro Cys Glu Asn Gly Gly Ile Cys Leu Pro
1 5 10 15
Gly Leu Ala Asp Gly Ser Phe Ser Cys Glu Cys Pro Asp Gly Phe Thr
20 25 30
Asp Pro Asn Cys Ser Ser Val Val Glu Val Ala Ser Asp Glu Glu Glu
35 40 45
Pro Thr Ser Ala Gly Pro Cys Thr Pro Asn Pro Cys His Asn Gly Gly
50 55 60
Thr Cys Glu Ile Ser Glu Ala Tyr Arg Gly Asp Thr Phe Ile Gly Tyr
65 70 75 80
Val Cys Lys Cys Pro Arg Gly Phe Asn Gly Ile His Cys Gln His Asn
85 90 95
Ile Asn Glu Cys Glu Val Glu Pro Cys Lys Asn Gly Gly Ile Cys Thr
100 105 110
Asp Leu Val Ala Asn Tyr Ser Cys Glu Cys Pro Gly Glu Phe Met Gly
115 120 125
Arg Asn Cys Gln Tyr Lys Asp Ala His Lys Ser Glu Val Ala His Arg
130 135 140
Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile Ala
145 150 155 160
Phe Ala Gln Tyr Leu Gln Gln Ser Pro Phe Glu Asp His Val Lys Leu
165 170 175
Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser
180 185 190
Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu
195 200 205
Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys
210 215 220
Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His Lys
225 230 235 240
Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp Val

245 250 255

Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr

260 265 270

Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu

275 280 285

Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln

290 295 300

Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg

305 310 315 320

Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser

325 330 335

Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg

340 345 350

Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu

355 360 365

Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp Leu

370 375 380

Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu

385 390 395 400

Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro

405 410 415

Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met

420 425 430

Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp

435 440 445

Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met Phe

450 455 460

Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu Leu

465 470 475 480

Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala

485 490 495

Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys
500 505 510
Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu
515 520 525
Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg
530 535 540
Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu Val
545 550 555 560
Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro Glu
565 570 575
Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn
580 585 590
Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr
595 600 605
Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser Ala
610 615 620
Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr
625 630 635 640
Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln
645 650 655
Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys
660 665 670
Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe
675 680 685
Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu
690 695 700
Glu Gly Lys Lys Leu Val Ala Cys Ser Glu Pro Leu Gly Met Lys Ser
705 710 715 720
Gly His Ile Gln Asp Tyr Gln Ile Thr Ala Ser Ser Ile Phe Arg Thr
725 730 735
Leu Asn Met Asp Met Phe Thr Trp Glu Pro Arg Lys Ala Arg Leu Asp

740 745 750

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

755 760 765

Gln Trp Leu Gln Val Asp Leu Leu Val Pro Thr Lys Val Thr Gly Ile

770 775 780

Ile Thr Gln Gly Ala Lys Asp Phe Gly His Val Gln Phe Val Gly Ser

785 790 795 800

Tyr Lys Leu Ala Tyr Ser Asn Asp Gly Glu His Trp Thr Val Tyr Gln

805 810 815

Asp Glu Lys Gln Arg Lys Asp Lys Val Phe Gln Gly Asn Phe Asp Asn

820 825 830

Asp Thr His Arg Lys Asn Val Ile Asp Pro Pro Ile Tyr Ala Arg His

835 840 845

Ile Arg Ile Leu Pro Trp Ser Trp Tyr Gly Arg Ile Thr Leu Arg Ser

850 855 860

Glu Leu Leu Gly Cys

865

<210> 136

<211> 2607

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 136

gacatctgcg accccaatcc ttgcgagaat ggcggcattt gtctgcctgg actggccgat 60

ggcagcttct cttgtgaatg ccccgatggc ttcacagacc ccaattgcag ctctgtggtg 120

gaagtggcca gcgacgagga agaacctaca agcgctggcc cctgcacacc caatccatgt 180

cataatggcg gaacctgcga gatcagcgag gcctacagag gcgatacctt catcggctac 240

gtgtgcaagt gccccagagg cttcaatggc atccactgcc agcacaacat caacgagtgc 300

gaggtggaac catgcaagaa cggcggcatc tgtaccgacc tggtggccaa ttactcttgc 360

gagtgccctg gcgagttcat gggcagaaac tgccagtaca aggacgccca caagagcgag 420

gtggcccaca gattcaagga cctgggcgaa gagaacttca aggccctggt gctgatcgcc 480

ttcgctcagt atctccagca gagcccttтc gaggaccacg tgaagctggt caacgaagtg 540

accgagttcg ccaagacctg tgtggccgat gagagcgccg agaactgtga caagagcctg 600

cacacactgt tcggcgacaa gctgtgtacc gtggccacac tgagagaaac ctacggcgag 660

atggccgact gctgtgccaa gcaagagccc gagagaaacg agtgcttcct ccagcacaag 720

gatgacaacc ccaacctgcc tagactcgtg cggcctgaag tggatgtgat gtgcaccgcc 780

tttcacgaca acgaggaaac cttcctgaag aagtacctgt acgagatcgc cagacggcac 840

ccctactttt atgcccctga gctgctgttc ttcgccaagc ggtataaggc cgccttcacc 900

gaatgttgcc aggccgctga taaggctgcc tgtctgctgc ctaagctgga cgagctgaga 960

gatgagggca aagccagctc tgccaagcag agactgaaat gcgccagcct ccagaagttc 1020

ggcgagagag cttttaaggc ctgggccgtt gccagactga gccagagatt tcctaaggcc 1080

gagtttgccg aggtgtccaa gctcgtgacc gatctgacaa aggtgcacac cgagtgctgt 1140

cacggcgatc tgctggaatg tgccgacgat agagccgacc tggccaagta tatctgcgag 1200

aaccaggaca gcatcagcag caagctgaaa gagtgctgcg agaagcccct gctggaaaag 1260

tctcactgta tcgccgaagt ggaaaacgac gagatgcccg ccgatctgcc ttctctggct 1320

gccgatttcg tggaaagcaa ggatgtgtgc aagaactacg ccgaggccaa agatgtgttt 1380

ctgggcatgt ttctgtatga gtacgcccgc agacaccccg actattctgt ggttctgctg 1440

ctgcggctgg ccaagacata cgagacaacc ctggaaaaat gctgcgccgc tgccgatcct 1500

cacgagtgtt atgccaaggt gttcgacgag ttcaagccac tggtggaaga accccagaac 1560

ctgatcaagc agaactgcga gctgttcgag cagctgggcg agtacaagtt ccagaatgcc 1620

ctgctcgtgc ggtacaccaa gaaagtgcct caggtgtcca cacctacact ggttgaggtg 1680

tcccggaatc tgggcaaagt gggcagcaag tgttgcaagc accctgaggc caagagaatg 1740

ccttgcgccg aggattacct gagcgtggtg ctgaatcagc tgtgcgtgct gcacgagaaa 1800

acccctgtgt ccgacagagt gaccaagtgc tgtaccgaga gcctcgtgaa cagaaggcct 1860

tgctttagcg ccctggaagt ggacgagaca tacgtgccca aagagttcaa cgccgagaca 1920

ttcaccttcc acgccgatat ctgcaccctg tccgagaaag agcggcagat caagaagcag 1980

acagccctgg tcgagctggt taagcacaag cccaaggcca ccaaagaaca gctgaaggcc 2040

gtgatggacg acttcgccgc ctttgtcgag aagtgctgca aggccgacga caaagagaca 2100

tgcttcgccg aagagggcaa gaaactggtg gcctgttctg agccactggg catgaagtct 2160

ggccacatcc aggattacca gatcaccgcc agcagcatct tcagaaccct gaacatggat 2220

atgttcacct gggagccccg gaaggccaga ctggataagc agggaaaagt gaacgcctgg 2280

accagcggcc acaatgacca gtctcagtgg ctgcaagtgg acctgctggt gcctaccaaa 2340

gtgaccggca tcatcacaca gggcgcaaag gatttcggcc acgtgcagtt tgtgggcagc 2400

tacaagctgg cctacagcaa cgatggcgag cactggacag tgtaccagga cgagaagcag 2460

cggaaggata aggtgttcca gggcaacttc gacaacgaca cccaccggaa gaacgtgatc 2520

gaccctccta tctacgcccg gcacatcaga atcctgcctt ggtcttggta cggccggatc 2580

accctgagaa gcgagctgct tggatgt 2607

<210> 137

<211> 781

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 137

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala
35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn
50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser
65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala
85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu
100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu
115 120 125

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

370 375 380

Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu

385 390 395 400

Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr

405 410 415

Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn

420 425 430

Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys

435 440 445

Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val

450 455 460

Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly

465 470 475 480

Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu

485 490 495

Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys

500 505 510

Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val

515 520 525

Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val

530 535 540

Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys

545 550 555 560

Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val

565 570 575

Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala

580 585 590

Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp

595 600 605

Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Cys

610 615 620

Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln Ile
625 630 635 640
Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser Trp
645 650 655
Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala Trp
660 665 670
Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu Gly
675 680 685
Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn Phe
690 695 700
Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn Asp
705 710 715 720
Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser Lys
725 730 735
Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu Phe
740 745 750
Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala Trp
755 760 765
His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys
770 775 780

<210> 138
<211> 2343
<212> ADN
<213> Trình tự Nhân tạo
<220>
<221> nguồn
<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"
<400> 138

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60
gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120
aaccactgcg agacaaagga tgcccacaag agcgaggtgg cccacagatt caaggatctg 180

ggcgaagaga acttcaaggc cctggtgctg atcgccttcg ctcagtatct ccagcagagc 240

cctttcgagg accacgtgaa gctggtcaac gaagtgaccg agttcgccaa gacctgtgtg 300

gccgatgaga gcgccgagaa ctgtgataag agcctgcaca ccctgttcgg cgacaagctg 360

tgtacagtgg ccacactgag agaaacctac ggcgagatgg ccgactgctg tgccaagcaa 420

gagcccgaga gaaacgagtg cttcctccag cacaaggacg acaaccccaa cctgcctaga 480

ctcgtgcgac ccgaagtgga tgtgatgtgc accgcctttc acgacaacga ggaaaccttc 540

ctgaagaagt acctgtacga gatcgccaga cggcacccct acttttatgc ccctgagctg 600

ctgttcttcg ccaagcggta taaggccgcc ttcaccgaat gttgccaggc cgctgataag 660

gctgcctgtc tgctgcctaa gctggacgag ctgagagatg agggcaaagc cagctctgcc 720

aagcagagac tgaaatgcgc cagcctccag aagttcggcg agagagcttt taaggcctgg 780

gccgttgcca gactgagcca gagatttcct aaggccgagt ttgccgaggt gtccaagctc 840

gtgaccgatc tgacaaaggt gcacaccgag tgctgtcacg gcgatctgct ggaatgtgcc 900

gacgatagag ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 960

ctgaaagagt gctgcgagaa gcccctgctg gaaaagtctc actgtatcgc cgaggtggaa 1020

aacgacgaga tgcctgccga tctgcctagc ctggctgccg atttcgtgga aagcaaggac 1080

gtgtgcaaga actacgccga ggccaaggat gtgtttctgg gcatgtttct gtatgagtac 1140

gcccgcagac accccgacta ttctgtggtt ctgctgctgc ggctggccaa aacctacgag 1200

acaaccctgg aaaaatgctg cgccgctgcc gatcctcacg agtgttatgc caaggtgttc 1260

gacgagttca agcctctggt ggaagaaccc cagaacctga tcaagcagaa ctgcgagctg 1320

ttcgagcagc tgggcgagta caagttccag aatgccctgc tcgtgcggta caccaagaaa 1380

gtgcctcagg tgtccacacc tacactggtt gaggtgtccc ggaatctggg caaagtgggc 1440

agcaagtgtt gcaagcaccc tgaggccaag agaatgcctt gcgccgagga ttacctgagc 1500

gtggtgctga atcagctgtg cgtgctgcac gagaaaaccc ctgtgtccga cagagtgacc 1560

aagtgctgta ccgagagcct cgtgaacaga aggccttgct ttagcgccct ggaagtggac 1620

gagacatacg tgcccaaaga gttcaacgcc gagacattca ccttccacgc cgacatctgc 1680

accctgtccg agaaagagcg gcagatcaag aagcagacag ccctggtcga gctggttaag 1740

cacaagccca aggccaccaa agaacagctg aaggccgtga tggacgactt cgccgccttt 1800

gtcgagaagt gctgcaaggc cgacgacaaa gagacatgct tcgccgaaga gggcaagaaa 1860

ctggtggcct gtgctaaccc tctgggcctg aagaacaaca gcatccccga taagcagatc 1920

accgccagca gcagctataa gacatggggc ctgcacctgt tcagctggaa cccttcttac 1980

gccagactgg acaagcaggg caacttcaat gcttgggtgg ccggcagcta cggcaatgat 2040

cagtggctgc aagtggacct gggcagcagc aaagaagtga caggcatcat cacccagggc 2100

gcaagaaatt tcggcagcgt gcagttcgtg gccagctaca aggtggccta cagcaacgat 2160

agcgccaact ggaccgagta tcaggaccct agaaccggca gctccaagat cttccccggc 2220

aactgggaca accacagcca caagaagaat ctgttcgaga cacccatcct ggccagatac 2280

gtgcggattc tgcctgtggc ctggcacaac agaatcgccc tgagactgga actgctgggc 2340

tgt 2343

<210> 139

<211> 623

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 139

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu

1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys

20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Asp Ala

35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn

50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Ser

65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala

85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu

100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu

115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg

130 135 140

Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg

145 150 155 160

Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn

165 170 175

Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His

180 185 190

Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys

195 200 205

Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu

210 215 220

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

225 230 235 240

Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala

245 250 255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala

260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His

275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala

290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys

305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile

325 330 335

Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala

340 345 350

Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala

355 360 365

Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His

370 375 380

Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu
385 390 395 400

Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr
405 410 415

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

Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys
435 440 445

Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val
450 455 460

Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly
465 470 475 480

Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu
485 490 495

Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys
500 505 510

Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val
515 520 525

Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val
530 535 540

Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys
545 550 555 560

Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val
565 570 575

Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala
580 585 590

Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp
595 600 605

Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala
610 615 620

<210> 140

<211> 1869

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 140

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaagga tgcccacaag agcgaggtgg cccacagatt caaggatctg 180

ggcgaagaga acttcaaggc cctggtgctg atcgccttcg ctcagtatct ccagcagagc 240

cctttcgagg accacgtgaa gctggtcaac gaagtgaccg agttcgccaa gacctgtgtg 300

gccgatgaga gcgccgagaa ctgtgataag agcctgcaca ccctgttcgg cgacaagctg 360

tgtacagtgg ccacactgag agaaacctac ggcgagatgg ccgactgctg tgccaagcaa 420

gagcccgaga gaaacgagtg cttcctccag cacaaggacg acaaccccaa cctgcctaga 480

ctcgtgcgac ccgaagtgga tgtgatgtgc accgcctttc acgacaacga ggaaaccttc 540

ctgaagaagt acctgtacga gatcgccaga cggcacccct acttttatgc ccctgagctg 600

ctgttcttcg ccaagcggta taaggccgcc ttcaccgaat gttgccaggc cgctgataag 660

gctgcctgtc tgctgcctaa gctggacgag ctgagagatg agggcaaagc cagctctgcc 720

aagcagagac tgaaatgcgc cagcctccag aagttcggcg agagagcttt taaggcctgg 780

gccgttgcca gactgagcca gagatttcct aaggccgagt ttgccgaggt gtccaagctc 840

gtgaccgatc tgacaaaggt gcacaccgag tgctgtcacg gcgatctgct ggaatgtgcc 900

gacgatagag ccgacctggc caagtacatc tgcgagaacc aggacagcat cagcagcaag 960

ctgaaagagt gctgcgagaa gcccctgctg gaaaagtctc actgtatcgc cgaggtggaa 1020

aacgacgaga tgcctgccga tctgcctagc ctggctgccg atttcgtgga aagcaaggac 1080

gtgtgcaaga actacgccga ggccaaggat gtgtttctgg gcatgtttct gtatgagtac 1140

gcccgcagac accccgacta ttctgtggtt ctgctgctgc ggctggccaa aacctacgag 1200

acaaccctgg aaaaatgctg cgccgctgcc gatcctcacg agtgttatgc caaggtgttc 1260

gacgagttca agcctctggt ggaagaaccc cagaacctga tcaagcagaa ctgcgagctg 1320

ttcgagcagc tgggcgagta caagttccag aatgccctgc tcgtgcggta caccaagaaa 1380

gtgcctcagg tgtccacacc tacactggtt gaggtgtccc ggaatctggg caaagtgggc 1440

agcaagtgtt gcaagcaccc tgaggccaag agaatgcctt gcgccgagga ttacctgagc 1500

gtggtgctga atcagctgtg cgtgctgcac gagaaaaccc ctgtgtccga cagagtgacc 1560

aagtgctgta ccgagagcct cgtgaacaga aggccttgct ttagcgccct ggaagtggac 1620

gagacatacg tgcccaaaga gttcaacgcc gagacattca ccttccacgc cgacatctgc 1680

accctgtccg agaaagagcg gcagatcaag aagcagacag ccctggtcga gctggttaag 1740

cacaagccca aggccaccaa agaacagctg aaggccgtga tggacgactt cgccgccttt 1800

gtcgagaagt gctgcaaggc cgacgacaaa gagacatgct tcgccgaaga gggcaagaaa 1860

ctggtggct 1869

<210> 141

<211> 160

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 141

Cys Val Glu Pro Leu Gly Leu Glu Asn Gly Asn Ile Ala Asn Ser Gln

1 5 10 15

Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp

20 25 30

Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp

35 40 45

Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu

50 55 60

Arg Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu

65 70 75 80

Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn

85 90 95

Gly His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu

100 105 110

Phe Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu

115 120 125

Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His

130 135 140

Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly

145 150 155 160

<210> 142

<211> 160

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 142

Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn Ile Ala Asn Ser Gln

1 5 10 15

Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu Gly Leu Gln His Trp

20 25 30

Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly Met Val Asn Ala Trp

35 40 45

Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile Gln Val Asn Leu Leu

50 55 60

Arg Arg Met Trp Val Thr Gly Val Val Thr Gln Gly Ala Ser Arg Leu

65 70 75 80

Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val Ala Tyr Ser Leu Asn

85 90 95

Gly His Glu Phe Asp Phe Ile His Asp Val Asn Lys Lys His Lys Glu

100 105 110

Phe Val Gly Asn Trp Asn Lys Asn Ala Val His Val Asn Leu Phe Glu

115 120 125

Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr Pro Thr Ser Cys His

130 135 140

Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly Cys Glu Leu Asn Gly
145 150 155 160

<210> 143

<211> 158

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 143

Cys Ala Asn Pro Leu Gly Leu Lys Asn Asn Ser Ile Pro Asp Lys Gln
1 5 10 15

Ile Thr Ala Ser Ser Ser Tyr Lys Thr Trp Gly Leu His Leu Phe Ser
20 25 30

Trp Asn Pro Ser Tyr Ala Arg Leu Asp Lys Gln Gly Asn Phe Asn Ala
35 40 45

Trp Val Ala Gly Ser Tyr Gly Asn Asp Gln Trp Leu Gln Val Asp Leu
50 55 60

Gly Ser Ser Lys Glu Val Thr Gly Ile Ile Thr Gln Gly Ala Arg Asn
65 70 75 80

Phe Gly Ser Val Gln Phe Val Ala Ser Tyr Lys Val Ala Tyr Ser Asn
85 90 95

Asp Ser Ala Asn Trp Thr Glu Tyr Gln Asp Pro Arg Thr Gly Ser Ser
100 105 110

Lys Ile Phe Pro Gly Asn Trp Asp Asn His Ser His Lys Lys Asn Leu
115 120 125

Phe Glu Thr Pro Ile Leu Ala Arg Tyr Val Arg Ile Leu Pro Val Ala
130 135 140

Trp His Asn Arg Ile Ala Leu Arg Leu Glu Leu Leu Gly Cys
145 150 155

<210> 144

<211> 161

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 144

Cys Ser Gly Pro Leu Gly Ile Glu Gly Gly Ile Ile Ser Asn Gln Gln
1 5 10 15

Ile Thr Ala Ser Ser Thr His Arg Ala Leu Phe Gly Leu Gln Lys Trp
20 25 30

Tyr Pro Tyr Tyr Ala Arg Leu Asn Lys Lys Gly Leu Ile Asn Ala Trp
35 40 45

Thr Ala Ala Glu Asn Asp Arg Trp Pro Trp Ile Gln Ile Asn Leu Gln
50 55 60

Arg Lys Met Arg Val Thr Gly Val Ile Thr Gln Gly Ala Lys Arg Ile
65 70 75 80

Gly Ser Pro Glu Tyr Ile Lys Ser Tyr Lys Ile Ala Tyr Ser Asn Asp
85 90 95

Gly Lys Thr Trp Ala Met Tyr Lys Val Lys Gly Thr Asn Glu Asp Met
100 105 110

Val Phe Arg Gly Asn Ile Asp Asn Asn Thr Pro Tyr Ala Asn Ser Phe
115 120 125

Thr Pro Pro Ile Lys Ala Gln Tyr Val Arg Leu Tyr Pro Gln Val Cys
130 135 140

Arg Arg His Cys Thr Leu Arg Met Glu Leu Leu Gly Cys Glu Leu Ser
145 150 155 160

Gly

<210> 145

<211> 162

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 145

Cys Ser Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr Gln
1 5 10 15

Ile Thr Ala Ser Ser Ile Phe Arg Thr Leu Asn Met Asp Met Phe Thr
20 25 30

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

Trp Thr Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp Leu
50 55 60

Leu Val Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys Asp
65 70 75 80

Phe Gly His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser Asn
85 90 95

Asp Gly Glu His Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys Asp
100 105 110

Lys Val Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn Val
115 120 125

Ile Asp Pro Pro Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp Ser
130 135 140

Trp Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys Thr Glu
145 150 155 160

Glu Glu

<210> 146

<211> 158

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 146

Cys Ser Glu Pro Leu Gly Met Lys Ser Gly His Ile Gln Asp Tyr Gln
1 5 10 15

Ile Thr Ala Ser Ser Ile Phe Arg Thr Leu Asn Met Asp Met Phe Thr
20 25 30

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

Trp Thr Ser Gly His Asn Asp Gln Ser Gln Trp Leu Gln Val Asp Leu
50 55 60

Leu Val Pro Thr Lys Val Thr Gly Ile Ile Thr Gln Gly Ala Lys Asp
65 70 75 80

Phe Gly His Val Gln Phe Val Gly Ser Tyr Lys Leu Ala Tyr Ser Asn
85 90 95

Asp Gly Glu His Trp Thr Val Tyr Gln Asp Glu Lys Gln Arg Lys Asp
100 105 110

Lys Val Phe Gln Gly Asn Phe Asp Asn Asp Thr His Arg Lys Asn Val
115 120 125

Ile Asp Pro Pro Ile Tyr Ala Arg His Ile Arg Ile Leu Pro Trp Ser
130 135 140

Trp Tyr Gly Arg Ile Thr Leu Arg Ser Glu Leu Leu Gly Cys
145 150 155

<210> 147

<211> 805

<212> PRT

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polypeptit tổng hợp"

<400> 147

Leu Asp Ile Cys Ser Lys Asn Pro Cys His Asn Gly Gly Leu Cys Glu
1 5 10 15

Glu Ile Ser Gln Glu Val Arg Gly Asp Val Phe Pro Ser Tyr Thr Cys
20 25 30

Thr Cys Leu Lys Gly Tyr Ala Gly Asn His Cys Glu Thr Lys Gly Ser
35 40 45

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
50 55 60

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
65 70 75 80

Gln Ser Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
85 90 95

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
100 105 110

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
115 120 125

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
130 135 140

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
145 150 155 160

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
165 170 175

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
180 185 190

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
195 200 205

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
210 215 220

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser

225 230 235 240

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu

245 250 255

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro

260 265 270

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

275 280 285

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp

290 295 300

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser

305 310 315 320

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His

325 330 335

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser

340 345 350

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

355 360 365

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg

370 375 380

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr

385 390 395 400

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu

405 410 415

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro

420 425 430

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu

435 440 445

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro

450 455 460

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys

465 470 475 480

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
485 490 495
Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
500 505 510
Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
515 520 525
Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
530 535 540
Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
545 550 555 560
Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
565 570 575
Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
580 585 590
Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
595 600 605
Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
610 615 620
Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly
625 630 635 640
Gly Ser Gly Gly Ser Cys Val Glu Pro Leu Gly Met Glu Asn Gly Asn
645 650 655
Ile Ala Asn Ser Gln Ile Ala Ala Ser Ser Val Arg Val Thr Phe Leu
660 665 670
Gly Leu Gln His Trp Val Pro Glu Leu Ala Arg Leu Asn Arg Ala Gly
675 680 685
Met Val Asn Ala Trp Thr Pro Ser Ser Asn Asp Asp Asn Pro Trp Ile
690 695 700
Gln Val Asn Leu Leu Arg Arg Met Trp Val Thr Gly Val Val Thr Gln
705 710 715 720
Gly Ala Ser Arg Leu Ala Ser His Glu Tyr Leu Lys Ala Phe Lys Val

725 730 735

Ala Tyr Ser Leu Asn Gly His Glu Phe Asp Phe Ile His Asp Val Asn

740 745 750

Lys Lys His Lys Glu Phe Val Gly Asn Trp Asn Lys Asn Ala Val His

755 760 765

Val Asn Leu Phe Glu Thr Pro Val Glu Ala Gln Tyr Val Arg Leu Tyr

770 775 780

Pro Thr Ser Cys His Thr Ala Cys Thr Leu Arg Phe Glu Leu Leu Gly

785 790 795 800

Cys Glu Leu Asn Gly

805

<210> 148

<211> 2415

<212> ADN

<213> Trình tự Nhân tạo

<220>

<221> nguồn

<223> /lưu ý="Mô tả trình tự nhân tạo: Polynucleotit tổng hợp"

<400> 148

ctggacatct gtagcaagaa cccttgccac aacggcggcc tgtgcgaaga gatttctcaa 60

gaagtgcggg gcgacgtttt ccccagctac acctgtacat gtctgaaggg ctacgccggc 120

aaccactgcg agacaaaagg atctgatgcc cacaagagcg aggtggccca cagattcaag 180

gatctgggcg aagagaactt caaggccctg gtgctgatcg ccttcgctca gtatctccag 240

cagagccctt tcgaggacca cgtgaagctg gtcaacgaag tgaccgagtt cgccaagacc 300

tgtgtggccg atgagagcgc cgagaactgt gataagagcc tgcacaccct gttcggcgac 360

aagctgtgta cagtggccac actgagagaa acctacggcg agatggccga ctgctgtgcc 420

aagcaagagc ccgagagaaa cgagtgcttc ctccagcaca aggacgacaa ccccaacctg 480

cctagactcg tgcgacccga agtggatgtg atgtgcaccg cctttcacga caacgaggaa 540

accttcctga agaagtacct gtacgagatc gccagacggc acccctactt ttatgccccc 600

gagctgctgt tcttcgccaa gcggtataag gccgccttca ccgaatgttg ccaggccgct 660

gataaggctg cctgtctgct gcctaagctg gacgagctga gagatgaggg caaagccagc 720

tctgccaagc agagactgaa atgcgccagc ctccagaagt tcggcgagag agcttttaag 780

gcctgggccg ttgccagact gagccagaga tttcctaagg ccgagtttgc cgaggtgtcc 840

aagctcgtga ccgatctgac aaaggtgcac accgagtgct gtcacggcga tctgctggaa 900

tgtgccgacg atagagccga cctggccaag tacatctgcg agaaccagga cagcatcagc 960

agcaagctga aagagtgctg cgagaagccc ctgctggaaa agtctcactg tatcgccgag 1020

gtggaaaacg acgagatgcc tgccgatctg cctagcctgg ctgccgattt cgtggaaagc 1080

aaggacgtgt gcaagaacta cgccgaggcc aaggatgtgt ttctgggcat gtttctgtat 1140

gagtacgccc gcagacaccc cgactattct gtggttctgc tgctgcggct ggccaaaacc 1200

tacgagacaa ccctggaaaa atgctgcgcc gctgccgatc ctcacgagtg ttatgccaag 1260

gtgttcgacg agttcaagcc tctggtggaa gaaccccaga acctgatcaa gcagaactgc 1320

gagctgttcg agcagctggg cgagtacaag ttccagaatg ccctgctcgt gcggtacacc 1380

aagaaagtgc ctcaggtgtc cacacctaca ctggttgagg tgtcccggaa tctgggcaaa 1440

gtgggcagca agtgttgcaa gcaccctgag gccaagagaa tgccttgcgc cgaggattac 1500

ctgagcgtgg tgctgaatca gctgtgcgtg ctgcacgaga aaacccctgt gtccgacaga 1560

gtgaccaagt gctgtaccga gagcctcgtg aacagaaggc cttgctttag cgccctggaa 1620

gtggacgaga catacgtgcc caaagagttc aacgccgaga cattcacctt ccacgccgac 1680

atctgcaccc tgtccgagaa agagcggcag atcaagaagc agacagccct ggtcgagctg 1740

gttaagcaca agcccaaggc caccaaagaa cagctgaagg ccgtgatgga cgacttcgcc 1800

gcctttgtcg agaagtgctg caaggccgac gacaaagaga catgcttcgc cgaagagggc 1860

aagaaactgg tggctgcctc tcaggctgct ctcggacttg gtggaagcgg aggaagtggt 1920

ggatctggcg gatcttgtgt ggaacccctc ggcatggaaa acggcaatat cgccaatagc 1980

cagattgccg ccagcagcgt cagagtgaca tttctgggac tgcaacactg ggtgcccgag 2040

ctggctagac tgaatagagc cggcatggtc aacgcctgga cacccagcag caacgacgat 2100

aatccctgga ttcaagtgaa cctgctgcgg cgtatgtggg tcacaggtgt tgttacacag 2160

ggcgcaagca gactggccag ccacgagtat ctgaaggcct ttaaggtggc ctacagcctg 2220

aacggccacg agttcgactt catccacgac gtgaacaaga agcacaaaga gtttgtcggc 2280

aactggaaca agaacgccgt gcacgtgaac ctgttcgaga cacctgtgga agcccagtac 2340

gtgcggctgt accctacaag ctgtcacacc gcctgcactc tgagattcga actgctggga 2400

tgcgagctga acggc 2415