



(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)



1-0052733

(51)^{2021.01} C07K 14/34; C12N 1/19; C12N 1/21; (13) B
C12P 13/00; C12N 15/53; C12N 15/63;
C12N 5/10; C12N 9/02; C12N 1/15;
C12N 15/31

(21) 1-2022-03555

(22) 06/11/2020

(86) PCT/JP2020/041581 06/11/2020

(87) WO/2021/090925 14/05/2021

(30) 2019-203523 08/11/2019 JP; 2019-233485 24/12/2019 JP; 2019-233484 24/12/2019 JP

(45) 27/10/2025 451

(43) 25/10/2022 415A

(73) KAO CORPORATION (JP)

14-10, Nihonbashi Kayabacho 1-chome, Chuo-ku, Tokyo 103-8210 Japan

(72) NONAKA, Kyoshiro (JP); TAKAHASHI, Fumikazu (JP).

(74) Công ty Luật TNHH T&G (TGVN)

(54) POLYPEPTIT CÓ HOẠT TÍNH HYDROXYL HÓA AXIT 4-AMINOBENZOIC,
PHƯƠNG PHÁP SẢN XUẤT POLYPEPTIT NÀY VÀ PHƯƠNG PHÁP CẢI
THIỆN HOẠT TÍNH HYDROXYL HÓA AXIT 4-AMINOBENZOIC

(21) 1-2022-03555

(57) Sáng chế đề cập đến polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic vượt trội và phương pháp để sử dụng polypeptit này. Sáng chế đề xuất polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó là leuxin.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic và việc sử dụng polypeptit này.

Tình trạng kỹ thuật của sáng chế

Polybenzoxazole (PBO) được biết đến là loại nhựa kỹ thuật có khả năng chịu nhiệt và độ bền cơ học vượt trội, và được sử dụng trong các vật liệu sợi và màng cách điện của các thiết bị bán dẫn, v.v. (Tài liệu phi sáng chế 1).

Khung benzoxazole được tạo thành do sự ngưng tụ của khung o-aminophenol với axit cacboxylic. Do đó, người ta kỳ vọng rằng các axit 4-amino-3-hydroxybenzoic (các 4,3-AHBA) có các nhóm chức này trong các phân tử của chúng sẽ hữu ích làm các monome PBO. Trên thực tế, việc tổng hợp và đánh giá tính chất vật lý của polybenzoxazole sử dụng 4,3-AHBA đã được nghiên cứu (Tài liệu phi sáng chế 2).

Các phương pháp sản xuất các hợp chất bằng cách lên men vi sinh vật sử dụng các tài nguyên tái tạo làm nguyên liệu ban đầu đã được chú ý trong những năm gần đây để giảm tải môi trường toàn cầu, v.v.. Ví dụ, việc sản xuất vi sinh vật và polyme hóa axit 3-amino-4-hydroxybenzoic (3,4-AHBA), có cấu trúc tương tự như 4,3-AHBA, đã được nghiên cứu (Tài liệu sáng chế 1).

Đối với việc sản xuất 4,3-AHBA, ví dụ, phương pháp tổng hợp để khử hóa học hợp chất thơm nitro đã được biết đến cho đến nay (Tài liệu sáng chế 2). Chiến lược khả thi cho phép sản xuất lên men 4,3-AHBA bằng phương pháp vi sinh là sự hydroxyl hóa ở vị trí 3 của axit 4-aminobenzoic (4-ABA) có thể được sinh tổng hợp trong vi sinh vật. Tuy nhiên, đối với phản ứng như vậy, đã có báo cáo đơn thuần rằng một số hydroxylaza axit 4-hydroxybenzoic có hoạt tính nhẹ (các Tài liệu phi sáng chế 3 và 4).

[Tài liệu sáng chế 1] JP-B-5445453

[Tài liệu sáng chế 2] JP-B-3821350

[Tài liệu phi sáng chế 1] Hiroki Murase, SENI GAKKAISHI (Seni To Kogyo, Journal of Fiber Science and Technology), Vol. 66, No. 6 (2010)

[Tài liệu phi sáng chế 2] Lon J. Mathias et al., Macromolecules, Vol. 18, No. 4, pp. 616-622 (1985)

[Tài liệu phi sáng chế 3] Barrie Entsch et al., The Journal of Biological Chemistry, Vol. 262, No. 13, pp. 6060-6068 (1987)

[Tài liệu phi sáng chế 4] Domenico L. Gatti et al., Biochemistry, Vol. 35, No. 2, pp. 567-578 (1996)

Bản chất kỹ thuật của sáng chế

Sáng chế đề cập đến các mục từ 1) đến 7) sau đây.

1) Polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, polypeptit này được chọn từ nhóm bao gồm A) đến C) sau đây:

A) polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó là leuxin;

B) polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 là phenylalanin; và

C) polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit

amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 là axit amin sau đây:

- (a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,
- (b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,
- (c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,
- (d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và
- (e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

2) Phương pháp sản xuất polypeptit đột biến có hoạt tính hydroxyl hóa axit 4-aminobenzoic, bao gồm bước thay thế gốc axit amin, bước thay thế này được chọn từ nhóm bao gồm A') đến C') sau đây:

A') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó bằng leuxin;

B') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 bằng phenylalanin; và

C') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 bằng axit amin sau đây:

- (a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,
- (b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,

(c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,

(d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và

(e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

3) Phương pháp cải thiện hoạt tính hydroxyl hóa axit 4-aminobenzoic, bao gồm bước thay thế gốc axit amin, bước thay thế này được chọn từ nhóm bao gồm A') đến C') sau đây:

A') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó bằng leuxin;

B') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 bằng phenylalanin; và

C') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 bằng axit amin sau đây:

(a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,

(b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,

(c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,

(d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và

(e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

4) Polynucleotit mã hóa polypeptit theo mục 1) hoặc 2).

- 5) Vector hoặc mảnh ADN bao gồm polynucleotit theo mục 4).
- 6) Tế bào được biến nạp bao gồm vector hoặc mảnh ADN theo mục 5).
- 7) Phương pháp sản xuất axit 4-amino-3-hydroxybenzoic, bao gồm bước nuôi cấy tế bào được biến nạp theo mục 6).

Mô tả chi tiết sáng chế

Sáng chế đề cập đến việc cung cấp polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic vượt trội và phương pháp sử dụng polypeptit này.

Các tác giả sáng chế đã phát hiện ra rằng thể đột biến axit 4-hydroxybenzoic hydroxylaza có trình tự axit amin cụ thể có hoạt tính hydroxyl hóa axit 4-aminobenzoic vượt trội và hữu ích trong việc sản xuất axit 4-amino-3-hydroxybenzoic.

Vì polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic theo sáng chế có hoạt tính hydroxyl hóa axit 4-aminobenzoic vượt trội, việc sử dụng chúng cho phép axit 4-amino-3-hydroxybenzoic được sản xuất một cách hiệu quả từ axit 4-aminobenzoic.

Trong bản mô tả này, độ tương đồng của trình tự axit amin hoặc trình tự nucleotit được tính bằng phương pháp Lipman-Pearson (Science, 1985, 227: 1435-1441). Cụ thể, độ tương đồng được tính bằng phân tích với kích thước đơn vị để so sánh (ktup) được đặt thành 2 bằng cách sử dụng chương trình phân tích tương đồng (tìm kiếm tương đồng) của phần mềm xử lý thông tin di truyền GENETYX Ver. 12.

Trong bản mô tả này, "vị trí tương ứng" trên trình tự axit amin hoặc trình tự nucleotit có thể được xác định bằng cách căn thẳng hàng trình tự quan tâm và trình tự tham chiếu (ví dụ, trình tự axit amin được thể hiện bởi SEQ ID NO: 2) để cung cấp sự tương đồng tối đa. Việc căn thẳng hàng các trình tự axit amin hoặc trình tự nucleotit có thể được thực hiện bằng cách sử dụng thuật toán đã biết trong lĩnh vực này, và các quy trình của chúng mà đã biết đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật. Việc căn thẳng hàng trình tự có thể được thực hiện, ví dụ, bằng cách sử dụng cài đặt mặc định của chương trình đa căn chỉnh Clustal W (Thompson, J.D. et al., 1994, Nucleic Acids Res. 22: 4673-4680). Ngoài ra, có thể sử dụng Clustal W2 hoặc Clustal

omega, phiên bản sửa đổi của Clustal W, Clustal W, Clustal W2 và Clustal omega có sẵn trên, ví dụ, trang web của Viện Tin sinh học châu Âu (EBI [www.ebi.ac.uk/index.html]) hoặc Ngân hàng dữ liệu ADN của Nhật Bản (DDBJ [www.ddbj.nig.ac.jp/searches-j.html]) do National Institute of Genetics (Viện Di truyền Quốc gia) điều hành. Vị trí của trình tự quan tâm được căn thẳng hàng với vị trí tùy ý của trình tự tham chiếu bằng cách căn thẳng hàng được đề cập ở trên được coi là "vị trí tương ứng" với vị trí tùy ý.

Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này có thể tinh chỉnh thêm việc căn thẳng hàng các trình tự axit amin thu được như đã mô tả ở trên để tối ưu hóa. Việc căn thẳng hàng tối ưu như vậy tốt hơn là được xác định xét đến sự giống nhau giữa các trình tự axit amin, tần suất của khoảng trống được chèn vào, v.v.. Trong bối cảnh này, sự giống nhau giữa các trình tự axit amin để chỉ tỷ lệ (%) của số lượng vị trí mà ở đó các gốc axit amin giống nhau hoặc tương tự có mặt trong cả hai trình tự với số lượng gốc axit amin có chiều dài đầy đủ khi hai trình tự axit amin này thẳng hàng. Các gốc axit amin tương tự có nghĩa là các gốc axit amin có tính chất tương tự nhau về độ phân cực hoặc điện tích và gây ra hiện tượng gọi là sự thay thế bảo tồn, trong số 20 axit amin cấu thành protein. Các nhóm bao gồm các gốc axit amin tương tự như vậy là đã biết rõ đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. Các ví dụ về chúng bao gồm, nhưng không giới hạn ở: arginin và lysin hoặc glutamin; axit glutamic và axit aspartic hoặc glutamin; serin và threonin hoặc alanin; glutamin và asparagin hoặc arginin; và leuxin và isoleuxin.

Trong bản mô tả này, "gốc axit amin" có nghĩa là bất kỳ trong số 20 gốc axit amin cấu thành protein, tức là, alanin (Ala hoặc A), arginin (Arg hoặc R), asparagin (Asn hoặc N), axit aspartic (Asp hoặc D), xystein (Cys hoặc C), glutamin (Gln hoặc Q), axit glutamic (Glu hoặc E), glyxin (Gly hoặc G), histidin (His hoặc H), isoleuxin (Ile hoặc I), leuxin (Leu hoặc L), lysin (Lys hoặc K), methionin (Met hoặc M), phenylalanin (Phe hoặc F), proline (Pro hoặc P), serin (Ser hoặc S), threonin (Thr hoặc T), tryptophan (Trp hoặc W), tyrosin (Tyr hoặc Y) và valin (Val hoặc V).

Trong bản mô tả này, "liên kết hoạt động" của gen với vùng điều khiển chẳng hạn như trình tự khởi động có nghĩa là gen được liên kết với vùng điều khiển sao cho gen có thể biểu hiện được dưới sự kiểm soát của vùng điều khiển. Các quy trình của "liên kết hoạt động" của gen với vùng điều khiển là đã biết rõ đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật.

Trong bản mô tả này, các thuật ngữ "ngược dòng" và "xuôi dòng" liên quan đến gen để chỉ quá trình ngược dòng và xuôi dòng theo hướng phiên mã của gen. Ví dụ, cụm từ "gen nằm xuôi dòng của trình tự khởi động" có nghĩa là gen nằm ở phía 3' của trình tự khởi động trong sợi có nghĩa ADN, và cụm từ "ngược dòng của gen" có nghĩa là vùng ở phía 5' của gen trong sợi có nghĩa ADN.

Trong bản mô tả này, thuật ngữ "gốc" được sử dụng cho chức năng, tính chất, hoặc đặc điểm của tế bào được sử dụng để chỉ ra rằng chức năng, tính chất, hoặc đặc điểm là cùng xuất xứ với tế bào. Ngược lại, thuật ngữ "ngoại lai" được sử dụng để chỉ ra rằng chức năng, tính chất, hoặc đặc điểm được đưa vào từ bên ngoài, không phải cùng xuất xứ với tế bào. Ví dụ, gen hoặc polynucleotit "ngoại lai" là gen hoặc polynucleotit được đưa từ bên ngoài vào tế bào. Gen hoặc polynucleotit ngoại lai có thể có nguồn gốc từ sinh vật cùng loài với tế bào chứa nó hoặc có thể có nguồn gốc từ sinh vật thuộc các loài khác với nó (tức là, gen hoặc polynucleotit khác loại).

<Polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic>

Polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic theo sáng chế (được gọi là "polypeptit theo sáng chế") được chọn từ nhóm bao gồm A) đến C) sau đây:

A) polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó là leuxin;

B) polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 là phenylalanin; và

C) polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 là axit amin sau:

- (a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,
- (b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,
- (c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,
- (d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và
- (e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

Polypeptit A) là polypeptit đột biến có hoạt tính hydroxyl hóa axit 4-aminobenzoic, bao gồm polypeptit đóng vai trò là tham chiếu, tức là, polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và trong đó gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó được thay thế bằng leuxin.

Polypeptit B) là polypeptit đột biến có hoạt tính hydroxyl hóa axit 4-aminobenzoic, bao gồm polypeptit đóng vai trò là tham chiếu, tức là, polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và trong đó gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 được thay thế bằng phenylalanin.

Polypeptit C) là polypeptit đột biến có hoạt tính hydroxyl hóa axit 4-aminobenzoic, bao gồm polypeptit đóng vai trò là tham chiếu, tức là, polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và trong đó gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí

47, 72, 210, 294 hoặc 385 được thay thế bằng bất kỳ axit amin nào trong số các axit amin (a) đến (e) được nêu ở trên.

Trong sáng chế, "hoạt tính hydroxyl hóa axit 4-aminobenzoic" có nghĩa là hoạt tính xúc tác quá trình hydroxyl hóa của axit 4-aminobenzoic, tốt hơn là hoạt tính xúc tác quá trình hydroxyl hóa ở vị trí 3 của axit 4-aminobenzoic.

Hoạt tính hydroxyl hóa axit 4-aminobenzoic có thể được xác định, như được thể hiện trong các ví dụ được đề cập sau, bằng cách nuôi cấy vi sinh vật sản xuất polypeptit theo sáng chế, và đo lượng axit 4-amino-3-hydroxybenzoic được sản xuất bởi HPLC hoặc tương tự.

Polypeptit A) theo sáng chế có thể được sản xuất bằng cách thay thế gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó bằng leuxin, trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic.

Trong bối cảnh này, polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic là polypeptit "mẹ" của polypeptit A) theo sáng chế.

Polypeptit mẹ dùng để chỉ polypeptit tham chiếu trở thành polypeptit A) theo sáng chế bằng cách đưa đột biến xác định trước vào gốc axit amin.

Polypeptit B) theo sáng chế có thể được sản xuất bằng cách thay thế gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với vị trí 201 hoặc 222 bằng phenylalanin, trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic.

Trong bối cảnh này, polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic là polypeptit "mẹ" của polypeptit B) theo sáng chế.

Polypeptit mẹ dùng để chỉ polypeptit tham chiếu trở thành polypeptit B) theo sáng chế bằng cách đưa đột biến xác định trước vào gốc axit amin.

Polypeptit C) theo sáng chế có thể được sản xuất bằng cách thay thế gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 bằng axit amin sau, trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic:

- (a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,
- (b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,
- (c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,
- (d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và
- (e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

Trong bối cảnh này, polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic là polypeptit "mẹ" của polypeptit C) theo sáng chế.

Polypeptit mẹ dùng để chỉ polypeptit tham chiếu trở thành polypeptit C) theo sáng chế bằng cách đưa đột biến xác định trước vào gốc axit amin.

Theo sáng chế, HFM122, polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 (Trình tự tham chiếu NCBI: WP_010920262.1), được gọi là axit 4-hydroxybenzoic-3-monooxygenaza (EC1.14.13.2). Axit 4-hydroxybenzoic-3-monooxygenaza là enzym có hoạt tính xúc tác tăng tốc bất kỳ một hoặc cả hai phản ứng để sản xuất axit protocatechuic bằng cách hydroxyl hóa vị trí 3 của axit 4-hydroxybenzoic và phản ứng nghịch đảo của chúng, và là enzym xúc tác quá trình hydroxyl hóa của axit 4-hydroxybenzoic (axit 4-hydroxybenzoic hydroxylaza).

HFM122 này được phát hiện có hoạt tính hydroxyl hóa axit 4-aminobenzoic bởi chủ đơn (Đơn sáng chế Nhật Bản số 2018-171849).

<Polypeptit mẹ của polypeptit A)>

Trong polypeptit A), các ví dụ về polypeptit gồm trình tự axit amin có độ tương đồng ít nhất 47% với trình tự axit amin được thể hiện bởi SEQ ID NO: 2, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic bao gồm polypeptit gồm trình tự axit amin có độ tương đồng ít nhất 47%, cụ thể, có độ tương đồng 47% hoặc cao hơn, tốt hơn nữa là 50% hoặc cao hơn, tốt hơn nữa là 60% hoặc cao hơn, tốt hơn nữa là 70% hoặc cao hơn, tốt hơn nữa là 80% hoặc cao hơn, tốt hơn nữa là 90% hoặc cao hơn, tốt hơn nữa là 95% hoặc cao hơn, tốt hơn nữa là 96% hoặc cao hơn, tốt hơn nữa là 97% hoặc cao hơn, tốt hơn nữa là 98% hoặc cao hơn, tốt hơn nữa là 99% hoặc cao hơn với trình tự axit amin được thể hiện bởi SEQ ID NO: 2, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic. Các ví dụ cụ thể của chúng bao gồm HFM388 (SEQ ID NO: 4; tương đồng trình tự axit amin với SEQ ID NO: 2: 62%, Trình tự tham chiếu NCBI: WP_010976283.1), HFM339 (SEQ ID NO: 6; tương đồng trình tự axit amin với SEQ ID NO: 2: 61%, Trình tự tham chiếu NCBI: WP_011157287.1), HFM77 (SEQ ID NO: 8; tương đồng trình tự axit amin với SEQ ID NO: 2: 51%, Trình tự tham chiếu NCBI: WP_011089160.1), HFM737 (SEQ ID NO: 10; tương đồng trình tự axit amin với SEQ ID NO: 2: 51%, Trình tự tham chiếu NCBI: WP_011519894.1), và HFMss0-1 (SEQ ID NO: 12; tương đồng trình tự axit amin với SEQ ID NO: 2: 47%, Trình tự tham chiếu NCBI: WP_027494688.1). Trong số đó, HFM737 hoặc HFMss0-1 được ưu tiên từ quan điểm hoạt tính hydroxyl hóa axit 4-aminobenzoic được sở hữu bởi polypeptit theo sáng chế.

Các ví dụ được ưu tiên của polypeptit "mẹ" bao gồm polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2, cũng như polypeptit gồm trình tự axit amin có độ tương đồng 90% hoặc cao hơn, tốt hơn nữa là 95% hoặc cao hơn, tốt hơn nữa là 96% hoặc cao hơn, tốt hơn nữa là 98% hoặc cao hơn với trình tự axit amin được thể hiện bởi SEQ ID NO: 2 và có hoạt tính hydroxyl hóa axit 4-aminobenzoic. Các ví dụ khác về chúng bao gồm polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10 hoặc SEQ ID NO: 12, hoặc polypeptit gồm trình tự axit amin có độ tương đồng 90% hoặc cao hơn, tốt hơn là 95%

hoặc cao hơn, tốt hơn nữa là 96 % hoặc cao hơn, tốt hơn nữa là 98% hoặc cao hơn với mỗi trình tự này, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic.

Tốt hơn là polypeptit mẹ có gốc valin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó. Đối với polypeptit A) theo sáng chế, được ưu tiên hơn là thay thế valin ở vị trí 47 hoặc vị trí tương ứng với nó bằng leuxin.

<Polypeptit mẹ của polypeptit B)>

Trong polypeptit B), các ví dụ về polypeptit gồm trình tự axit amin có độ tương đồng ít nhất 51% với trình tự axit amin được thể hiện bởi SEQ ID NO: 2, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic bao gồm polypeptit gồm trình tự axit amin có độ tương đồng ít nhất 51%, cụ thể, độ tương đồng 51% hoặc cao hơn, tốt hơn là 60% hoặc cao hơn, tốt hơn nữa là 70% hoặc cao hơn, tốt hơn nữa là 80% hoặc cao hơn, tốt hơn nữa là 90% hoặc cao hơn, tốt hơn nữa là 95% hoặc cao hơn, tốt hơn nữa là 96% hoặc cao hơn, tốt hơn nữa là 97% hoặc cao hơn, tốt hơn nữa là 98% hoặc cao hơn, tốt hơn nữa là 99% hoặc cao hơn với trình tự axit amin được thể hiện bởi SEQ ID NO: 2, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic. Các ví dụ cụ thể của chúng bao gồm HFM388 (SEQ ID NO: 4; tương đồng trình tự axit amin với SEQ ID NO: 2: 62%, Trình tự tham chiếu NCBI: WP_010976283.1), HFM339 (SEQ ID NO: 6; tương đồng trình tự axit amin với SEQ ID NO: 2: 61%, Trình tự tham chiếu NCBI: WP_011157287.1), và HFM77 (SEQ ID NO: 8; tương đồng trình tự axit amin với SEQ ID NO: 2: 51%, Trình tự tham chiếu NCBI: WP_011089160.1). Trong số đó, HFM388 hoặc HFM339 được ưu tiên từ quan điểm hoạt tính hydroxyl hóa axit 4-aminobenzoic được sở hữu bởi polypeptit theo sáng chế.

Các ví dụ được ưu tiên về polypeptit "mẹ" bao gồm polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 cũng như polypeptit gồm trình tự axit amin có độ tương đồng 90% hoặc cao hơn, tốt hơn nữa là 95% hoặc cao hơn, tốt hơn nữa là 96% hoặc cao hơn, tốt hơn nữa là 98% hoặc cao hơn với trình tự axit amin được thể hiện bởi SEQ ID NO: 2, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic. Các ví dụ khác của chúng bao gồm polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID

NO: 4, SEQ ID NO: 6 hoặc SEQ ID NO: 8, hoặc polypeptit gồm trình tự axit amin có độ tương đồng 90% hoặc cao hơn, tốt hơn là 95% hoặc cao hơn, tốt hơn nữa là 96% hoặc cao hơn, tốt hơn nữa là 98% hoặc cao hơn với mỗi trình tự này, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic.

Polypeptit mẹ tốt hơn là có gốc tyrosin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với vị trí 201 hoặc 222. Polypeptit B) theo sáng chế tốt hơn nữa là polypeptit đột biến trong đó tyrosin ở vị trí 201 hoặc 222 hoặc vị trí tương ứng với vị trí 201 hoặc 222 được thay thế bằng phenylalanin. Đối với vị trí tương ứng với vị trí 201 hoặc 222 của SEQ ID NO: 2, ví dụ, các vị trí 201 và 222 trong SEQ ID NO: 4, các vị trí 201 và 222 trong SEQ ID NO: 6, và các vị trí 203 và 224 trong SEQ ID NO: 8 tương ứng với các vị trí này.

Do đó, tốt hơn là polypeptit mẹ có gốc tyrosin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 4 hoặc vị trí tương ứng với vị trí 201 hoặc 222. Polypeptit B) theo sáng chế tốt hơn nữa là polypeptit đột biến trong đó tyrosin ở vị trí 201 hoặc 222 hoặc vị trí tương ứng với vị trí 201 hoặc 222 được thay thế bằng phenylalanin.

Polypeptit mẹ tốt hơn là có gốc tyrosin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 6 hoặc vị trí tương ứng với vị trí 201 hoặc 222. Polypeptit B) theo sáng chế tốt hơn nữa là polypeptit đột biến trong đó tyrosin ở vị trí 201 hoặc 222 hoặc vị trí tương ứng với vị trí 201 hoặc 222 được thay thế bằng phenylalanin.

Polypeptit mẹ tốt hơn là có gốc tyrosin ở vị trí 203 hoặc 224 của trình tự axit amin được thể hiện bởi SEQ ID NO: 8 hoặc vị trí tương ứng với vị trí 203 hoặc 224. Polypeptit B) theo sáng chế tốt hơn nữa là polypeptit đột biến trong đó tyrosin ở vị trí 203 hoặc 224 hoặc vị trí tương ứng với vị trí 203 hoặc 224 được thay thế bằng phenylalanin.

<Polypeptit mẹ của polypeptit C)>

Trong polypeptit C), các ví dụ về polypeptit gồm trình tự axit amin có độ tương đồng ít nhất 90% với trình tự axit amin được thể hiện bởi SEQ ID NO: 2, và có hoạt

tính hydroxyl hóa axit 4-aminobenzoic bao gồm polypeptit gồm trình tự axit amin có độ tương đồng ít nhất 90%, cụ thể, độ tương đồng 90% hoặc cao hơn, tốt hơn là 95% hoặc cao hơn, tốt hơn nữa là 96% hoặc cao hơn, tốt hơn nữa là 97% hoặc cao hơn, tốt hơn nữa là 98% hoặc cao hơn, tốt hơn nữa là 99% hoặc cao hơn với trình tự axit amin được thể hiện bởi SEQ ID NO: 2, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic.

Polypeptit mẹ tốt hơn là có gốc valin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó, tốt hơn là có gốc histidin ở vị trí 72 của chúng hoặc vị trí tương ứng với nó, tốt hơn là có gốc leuxin ở vị trí 210 của chúng hoặc vị trí tương ứng với nó, tốt hơn là có gốc threonin ở vị trí 294 của chúng hoặc vị trí tương ứng với nó, hoặc tốt hơn là có gốc tyrosin ở vị trí 385 của chúng hoặc vị trí tương ứng với nó. Đối với polypeptit C) theo sáng chế, được ưu tiên hơn là thay thế valin ở vị trí 47 hoặc vị trí tương ứng với nó bằng isoleuxin, serin, threonin, xystein, methionin hoặc glutamin, thay thế histidin ở vị trí 72 hoặc vị trí tương ứng với nó bằng alanin hoặc methionin, thay thế leuxin ở vị trí 210 hoặc vị trí tương ứng với nó bằng methionin, thay thế threonin ở vị trí 294 hoặc vị trí tương ứng với nó bằng alanin, glyxin, xystein hoặc serin, hoặc thay thế tyrosin ở vị trí 385 hoặc vị trí tương ứng với nó bằng valin, leuxin, isoleuxin hoặc methionin, được ưu tiên hơn là thay thế threonin ở vị trí 294 hoặc vị trí tương ứng với nó bằng serin, thay thế valin ở vị trí 47 hoặc vị trí tương ứng với nó bằng isoleuxin, threonin, methionin hoặc glutamin, hoặc thay thế histidin ở vị trí 72 hoặc vị trí tương ứng với nó bằng methionin, và được ưu tiên hơn nữa là thay thế threonin ở vị trí 294 hoặc vị trí tương ứng với nó bằng serin hoặc thay thế valin ở vị trí 47 hoặc vị trí tương ứng với nó bằng isoleuxin.

<Polynucleotit mã hóa polypeptit theo sáng chế>

Theo sáng chế, bất kỳ kỹ thuật gây đột biến nào đã biết trong lĩnh vực này đều có thể được sử dụng như một phương pháp gây đột biến gốc axit amin của polypeptit mẹ. Ví dụ, trình tự nucleotit mã hóa gốc axit amin bị đột biến trong polynucleotit mã hóa trình tự axit amin của polypeptit mẹ (sau đây, còn được gọi là gen mẹ) có thể bị đột biến thành trình tự nucleotit mã hóa gốc axit amin bị đột biến để thu được polynucleotit mã hóa polypeptit theo sáng chế.

Về cơ bản, việc đưa đột biến quan tâm vào gen mẹ có thể được thực hiện bằng cách sử dụng phương pháp gây đột biến hướng vị trí khác nhau bất kỳ mà đã biết rõ đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. Ví dụ, phương pháp gây đột biến hướng vị trí có thể được thực hiện, bằng bất kỳ phương pháp nào chẳng hạn như PCR nghịch đảo hoặc gắn môi. Có thể sử dụng kit có sẵn trên thị trường để gây đột biến hướng vị trí (ví dụ, QuikChange II Site-Directed Mutagenesis Kit hoặc QuikChange Multi Site-Directed Mutagenesis Kit của Agilent Technologies, Inc.).

Sự gây đột biến hướng vị trí của gen mẹ có thể được thực hiện một cách tổng quát nhất bằng cách sử dụng các đoạn môi cho các đột biến chứa đột biến nucleotit được đưa vào. Các đoạn môi cho các đột biến có thể được thiết kế để chứa trình tự nucleotit được gắn vào vùng chứa trình tự nucleotit mã hóa gốc axit amin bị đột biến trong gen mẹ, và có trình tự nucleotit (codon) mã hóa gốc axit amin bị đột biến thay cho trình tự nucleotit (codon) mã hóa gốc axit amin bị đột biến. Các trình tự nucleotit (các codon) mã hóa các gốc axit amin không bị đột biến hoặc bị đột biến có thể được nhận biết bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này và lựa chọn một cách thích hợp dựa trên các tài liệu thông thường hoặc tương tự. Ngoài ra, việc đột biến điểm trực tiếp có thể sử dụng phương pháp liên kết các mảnh ADN được khuếch đại tương ứng từ phía ngược dòng và xuôi dòng của vị trí đột biến bằng cách sử dụng riêng rẽ hai đoạn môi bổ sung chứa đột biến nucleotit được đưa vào, trong một mảnh bằng kỹ thuật PCR SOE (ghép nối bằng sự kéo dài chồng chéo- splicing by overlap extension) (Gene, 1989, 77 (1): p. 61-68).

ADN khuôn chứa gen mẹ có thể được chuẩn bị bằng cách chiết ADN bộ gen theo phương pháp thông thường từ vi sinh vật sản xuất axit 4-hydroxybenzoic hydroxylaza được đề cập ở trên, hoặc bằng cách chiết ARN từ đó và tổng hợp cADN bằng cách phiên mã ngược. Ngoài ra, trình tự nucleotit tương ứng có thể được tổng hợp hóa học trên cơ sở trình tự axit amin của polypeptit mẹ, và được sử dụng làm ADN khuôn. Các trình tự ADN chứa các trình tự nucleotit mã hóa HFM122, HFM388, HFM339, HFM77, HFM737, và HFMss0-1 đã được đề cập là các polypeptit có hoạt

tính hydroxyl hóa axit 4-aminobenzoic nêu trong SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9 và SEQ ID NO: 11, tương ứng.

Các đoạn mồi cho các đột biến có thể được chuẩn bị bằng phương pháp tổng hợp oligonucleotit được nhiều người biết đến như phương pháp phosphoramidite (Nucleic Acids Research, 1989, 17: 7059-7071). Việc tổng hợp đoạn mồi như vậy có thể được thực hiện, ví dụ, bằng cách sử dụng thiết bị tổng hợp oligonucleotit có sẵn trên thị trường (do Applied Biosystems Inc. (ABI) sản xuất, v.v.). Polynucleotit mã hóa polypeptit theo sáng chế có đột biến quan tâm có thể thu được bằng cách gây đột biến hướng vào vị trí như đã mô tả ở trên với gen mẹ là ADN khuôn bằng cách sử dụng bộ mồi chứa các đoạn mồi cho các đột biến.

Polynucleotit mã hóa polypeptit theo sáng chế có thể chứa ADN sợi đơn hoặc sợi kép, cADN, ARN hoặc các axit nucleic nhân tạo khác. ADN, cADN và ARN có thể được tổng hợp hóa học. Polynucleotit cũng có thể chứa khung đọc mở (ORF) cũng như trình tự nucleotit của vùng không dịch mã (UTR). Polynucleotit có thể được tối ưu hóa codon cho thể biến nạp để sản xuất polypeptit đột biến theo sáng chế. Thông tin về các codon được sử dụng bởi các sinh vật khác nhau có sẵn từ Cơ sở dữ liệu sử dụng Codon ([\[www.kazusa.or.jp/codon/\]](http://www.kazusa.or.jp/codon/)).

<Vector hoặc mảnh ADN>

Polynucleotit thu được mã hóa polypeptit theo sáng chế có thể được chèn vào vector. Vector chứa polynucleotit là vector biểu hiện. Tốt hơn là, vector là vector biểu hiện có thể đưa polynucleotit mã hóa polypeptit theo sáng chế vào vi sinh vật chủ và cho phép polynucleotit được biểu hiện trong vi sinh vật chủ. Tốt hơn là, vector chứa polynucleotit mã hóa polypeptit theo sáng chế, và vùng điều khiển được liên kết hoạt động với nó. Vector có thể là vector có khả năng sinh sôi và sao chép tự động bên ngoài nhiễm sắc thể (tức là, trong plasmid), hoặc có thể là vector được hợp nhất vào nhiễm sắc thể.

Các ví dụ cụ thể về vector bao gồm pBluescript II SK(-) (Stratagene California), các vector thuộc hệ pUC chẳng hạn như pUC18/19 và pUC118/119 (Takara Bio Inc.), các vector thuộc hệ pET (Takara Bio Inc.), các vector thuộc hệ pGEX (GE Healthcare

Japan Corp.), các vector thuộc hệ pCold (Takara Bio Inc.), pHY300PLK (Takara Bio Inc.), pUB110 (Mckenzie, T. et al., 1986, Plasmid 15 (2): 93-103), pBR322 (Takara Bio Inc.), pRS403 (Stratagene California), pMW218/219 (Nippon Gene Co., Ltd.), Các vector thuộc hệ pRI chẳng hạn như pRI909/910 (Takara Bio Inc.), các vector thuộc hệ pBI (Clontech Laboratories, Inc.), các vector thuộc hệ IN3 (Inplanta Innovations Inc.), pPTR1/2 (Takara Bio Inc.), pDJB2 (D.J. Ballance et al., Gene, 36, 321-331, 1985), pAB4-1 (van Hartingsveldt W et al., Mol Gen Genet, 206, 71-75, 1987), pLeu4 (M.I.G. Roncero et al., Gene, 84, 335-343, 1989), pPyr225 (C.D. Skory et al., Mol Genet Genomics, 268, 397-406, 2002), và pFG1 (Gruber, F. et al., Curr Genet, 18, 447-451, 1990).

Polynucleotit mã hóa polypeptit theo sáng chế có thể được tạo cấu trúc như mảnh ADN chứa polynucleotit tương tự. Các ví dụ về mảnh ADN bao gồm các mảnh ADN được khuếch đại bằng PCR và các mảnh ADN được phân cắt bằng enzym giới hạn. Tốt hơn là, mảnh ADN có thể là cat-xet biểu hiện chứa polynucleotit mã hóa polypeptit theo sáng chế, và vùng điều khiển được liên kết hoạt động với nó.

Vùng điều khiển trong vector hoặc mảnh ADN là trình tự cho phép polynucleotit mã hóa polypeptit theo sáng chế được biểu hiện trong tế bào chủ trong đó vector hoặc mảnh ADN đã được đưa vào. Các ví dụ về chúng bao gồm các vùng điều chỉnh biểu hiện như các trình tự khởi động và trình tự kết thúc, và các điểm khởi đầu sao chép. Loại vùng điều khiển có thể được lựa chọn một cách thích hợp tùy theo loại vi sinh vật chủ trong đó vector hoặc mảnh ADN đã được đưa vào. Nếu cần, vector hoặc mảnh ADN có thể có thêm chất chỉ thị chọn lọc như gen kháng thuốc kháng sinh hoặc gen liên quan đến tổng hợp axit amin (ví dụ, gen kháng thuốc như ampicillin, neomycin, kanamycin, hoặc chloramphenicol).

Vector hoặc mảnh ADN có thể chứa trình tự polynucleotit mã hóa polypeptit cần thiết cho quá trình sinh tổng hợp axit 4-aminobenzoic. Các ví dụ về polypeptit cần thiết cho quá trình sinh tổng hợp axit 4-aminobenzoic bao gồm 4-amino-4-deoxychorismat synthaza (pabAB) và 4-amino-4-deoxychorismat lyaza (pabC).

Liên kết của polynucleotit mã hóa polypeptit theo sáng chế với vùng điều khiển hoặc trình tự gen chỉ thị có thể được thực hiện bằng phương pháp như SOE-PCR được đề cập ở trên. Các quy trình đưa trình tự gen vào vector đã được biết đến nhiều trong lĩnh vực kỹ thuật này. Loại vùng điều khiển như vùng trình tự khởi động, trình tự kết thúc, hoặc vùng tín hiệu tiết không bị giới hạn cụ thể, và trình tự khởi động hoặc trình tự tín hiệu tiết thường được sử dụng có thể được lựa chọn và sử dụng một cách thích hợp tùy theo vật chủ tiếp nhận.

Các ví dụ được ưu tiên về vùng điều khiển bao gồm, nhưng không chỉ giới hạn cụ thể ở, các vùng điều khiển mạnh có thể tăng cường sự biểu hiện so với loại đại, ví dụ, các trình tự khởi động biểu hiện cao đã được biết đến trong lĩnh vực này chẳng hạn như trình tự khởi động T7, trình tự khởi động lac, trình tự khởi động tac, và trình tự khởi động trp.

<Tế bào được biến nạp>

Tế bào được biến nạp theo sáng chế có thể thu được bằng cách đưa vector bao gồm polynucleotit mã hóa polypeptit theo sáng chế vào vật chủ, hoặc bằng cách đưa mảnh ADN chứa polynucleotit mã hóa polypeptit theo sáng chế vào hệ gen của vật chủ.

Tế bào được biến nạp là tế bào chứa polynucleotit mã hóa polypeptit theo sáng chế một cách rõ ràng, và có thể là tế bào có sự biểu hiện tăng cường của polynucleotit, và bằng cách mở rộng, là tế bào có sự biểu hiện tăng cường của polypeptit theo sáng chế.

Bất kỳ tế bào nào của nấm, men, xạ khuẩn, *E. coli*, *Bacillus subtilis*, hoặc các loại tương tự có thể được sử dụng làm tế bào chủ, và *E. coli* hoặc xạ khuẩn được ưu tiên. Các ví dụ về xạ khuẩn bao gồm vi khuẩn thuộc chi *Corynebacterium*, vi khuẩn thuộc chi *Mycovir khuẩn*, vi khuẩn thuộc chi *Rhodococcus*, vi khuẩn thuộc chi *Streptomyces*, và vi khuẩn thuộc chi *Propionivir khuẩn*. Vi khuẩn thuộc giống *Corynebacterium* được ưu tiên, và *Corynebacterium glutamicum* được ưu tiên hơn.

Trong số những loại khác, vi sinh vật có thể cung cấp axit 4-aminobenzoic đóng vai trò là các chất nền trong quá trình sinh tổng hợp axit 4-amino-3-hydroxybenzoic

được ưu tiên, và vi sinh vật có khả năng cung cấp axit 4-aminobenzoic tăng cường được ưu tiên hơn. Các ví dụ về phương pháp tăng cường khả năng cung cấp axit 4-aminobenzoic của vi sinh vật bao gồm phương pháp đưa vào vi sinh vật, vector chứa polynucleotit mã hóa polypeptit cần thiết cho quá trình sinh tổng hợp axit 4-aminobenzoic, và vùng điều khiển được liên kết hoạt động với nó, và phương pháp thay thế vùng điều khiển của polynucleotit mã hóa polypeptit cần thiết cho quá trình sinh tổng hợp axit 4-aminobenzoic, vốn ban đầu được mang bởi vi sinh vật, bằng trình tự khởi động biểu hiện mạnh.

Các ví dụ về phương pháp đưa vector hoặc mảnh ADN vào vật chủ bao gồm, ví dụ, tạo lỗ xung điện, biến nạp, chuyển nạp, liên hợp, phương pháp nguyên sinh chất, phương pháp súng bắn hạt hoặc phương pháp vi khuẩn *Agrobacterium*.

Các ví dụ về phương pháp đưa polynucleotit vào bộ gen của vật chủ bao gồm, nhưng không giới hạn cụ thể ở, phương pháp trao đổi chéo kép sử dụng mảnh ADN chứa polynucleotit. Đoạn ADN có thể được đưa vào xuôi dòng của trình tự khởi động của gen có mức độ biểu hiện lớn trong tế bào chủ được đề cập ở trên, hoặc một đoạn của mảnh ADN được liên kết một cách có thể hoạt động được với vùng điều khiển được đề cập ở trên có thể được chuẩn bị trước, và đoạn liên kết có thể được đưa vào bộ gen của vật chủ. Đoạn ADN có thể được liên kết thêm trước với một chất chỉ thị (gen kháng thuốc, gen bổ sung dưỡng chất, v.v..) để chọn một tế bào chứa polynucleotit theo sáng chế một cách thích hợp.

Tế bào được biến nạp chứa vector hoặc mảnh ADN quan tâm có thể được chọn bằng cách sử dụng chất chỉ thị chọn lọc. Ví dụ, khi chất chỉ thị chọn lọc là gen kháng thuốc kháng sinh, tế bào được biến nạp trong đó vector hoặc mảnh ADN quan tâm đã được đưa vào có thể được chọn bằng cách nuôi cấy trong môi trường nuôi cấy có bổ sung kháng sinh. Ví dụ, khi chất chỉ thị chọn lọc là gen liên quan đến tổng hợp axit amin, tế bào được biến nạp chứa vector hoặc mảnh ADN quan tâm có thể được chọn bằng cách sử dụng sự có mặt hoặc không có mặt của sinh vật khuyết dưỡng axit amin này làm chỉ số sau khi đưa gen vào vi sinh vật chủ cần axit amin này. Ngoài ra, việc

đưa vector hoặc mảnh ADN quan tâm vào có thể được xác nhận bằng cách kiểm tra trình tự ADN của tế bào được biến nạp bằng PCR hoặc tương tự.

Tế bào được biến nạp thu được ở trên được nuôi cấy trong môi trường nuôi cấy thích hợp để polynucleotit được đưa vào tế bào được biểu hiện để sản xuất polypeptit theo sáng chế. Cụ thể, tế bào được biến nạp có khả năng đóng vai trò như một vi khuẩn sản xuất polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic. Trong trường hợp nuôi cấy tế bào được biến nạp theo sáng chế, như được thể hiện trong các ví dụ được đề cập sau, năng suất của axit 4-amino-3-hydroxybenzoic được cải thiện so với trường hợp sử dụng tế bào được biến nạp sản xuất polypeptit mẹ.

Cụ thể, trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự đột biến để thay thế gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó bằng leuxin có ích trong việc cải thiện hoạt tính hydroxyl hóa axit 4-aminobenzoic và cũng có ích trong việc cải thiện năng suất của axit 4-amino-3-hydroxybenzoic.

Trong một polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự đột biến để thay thế gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với vị trí 201 hoặc 222 bằng phenylalanin có ích trong việc cải thiện hoạt tính hydroxyl hóa axit 4-aminobenzoic và cũng có ích trong việc cải thiện năng suất của axit 4-amino-3-hydroxybenzoic.

Trong một polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự đột biến để thay thế gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 bằng axit amin sau có ích trong việc cải thiện hoạt tính hydroxyl hóa axit 4-aminobenzoic và cũng có ích trong việc cải thiện năng suất của axit 4-amino-3-hydroxybenzoic:

(a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,

(b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,

(c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,

(d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và

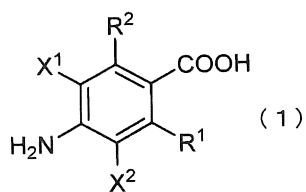
(e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

Hơn nữa, tế bào được biến nạp theo sáng chế đóng vai trò như một vi khuẩn sản xuất polypeptit và có hoạt tính hydroxyl hóa axit 4-aminobenzoic được cải thiện và đóng vai trò như một chủng có ích sản xuất axit 4-amino-3-hydroxybenzoic.

<Sản xuất axit 4-amino-3-hydroxybenzoic >

Phương pháp sản xuất axit 4-amino-3-hydroxybenzoic theo sáng chế bao gồm bước nuôi cấy tế bào được biến nạp theo sáng chế, và axit 4-amino-3-hydroxybenzoic có thể thu được bằng cách thu gom axit 4-amino-3-hydroxybenzoic từ môi trường nuôi cấy.

Theo sáng chế, các ví dụ cụ thể về axit 4-amino-3-hydroxybenzoic bao gồm các dẫn xuất axit 4-amino-3-hydroxybenzoic có công thức (1) sau đây:



trong đó R^1 là nguyên tử hydro, nhóm hydroxy (-OH), nhóm metoxy (-OCH₃), nhóm amin (-NH₂), nguyên tử flo (-F), nguyên tử clo (-Cl), nguyên tử brom (-Br), nguyên tử iốt (-I), nhóm cacboxy (-COOH), nhóm metyl (-CH₃), hoặc nhóm etyl (-CH₂CH₃), R^2 là nguyên tử hydro, nhóm hydroxy (-OH), nhóm metoxy (-OCH₃), nhóm amin (-NH₂), nguyên tử flo (-F), nguyên tử clo (-Cl), nguyên tử brom (-Br), nguyên tử iốt (-I), nhóm cacboxy (-COOH), nhóm metyl (-CH₃) hoặc nhóm etyl (-CH₂CH₃), và X^1 và X^2 mỗi loại là nguyên tử hydro hoặc nhóm hydroxy, ít nhất một trong số đó là nhóm hydroxy.

Nhóm chức R^1 tốt hơn là nguyên tử hydro, nhóm hydroxy (-OH), nhóm metoxy (-OCH₃), nguyên tử flo (-F) hoặc nhóm metyl (-CH₃).

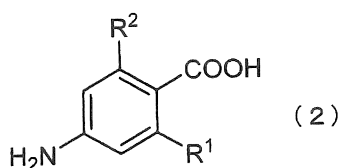
Nhóm chức R^2 tốt hơn là nguyên tử hydro, nhóm hydroxy (-OH), nhóm metoxy (-OCH₃), nguyên tử flo (-F) hoặc nhóm metyl (-CH₃).

Tốt hơn nữa là, cả R^1 và R^2 đều là nguyên tử hydro.

Cả X^1 và X^2 đều có thể là nhóm hydroxy, và một trong số X^1 hoặc X^2 tốt hơn là nhóm hydroxy.

Môi trường nuôi cấy có thể chứa, nếu cần, các axit 4-aminobenzoic đóng vai trò là cơ chất trong quá trình sinh tổng hợp các axit 4-amino-3-hydroxybenzoic.

Trong bối cảnh này, các ví dụ về axit 4-aminobenzoic bao gồm các dẫn xuất axit 4-aminobenzoic có công thức (2) sau đây:



trong đó R^1 và R^2 như được xác định ở trên.

Bất kỳ môi trường nuôi cấy tự nhiên và môi trường nuôi cấy tổng hợp nào cũng có thể được sử dụng làm môi trường nuôi cấy để nuôi cấy tế bào được biến nạp miễn là môi trường nuôi cấy chứa nguồn cacbon, nguồn nitơ, muối vô cơ, v.v. và cho phép nuôi cấy hiệu quả tế bào được biến nạp theo sáng chế. Ví dụ, các sacarit như glucoza, các polyol như glyxerin, các rượu như etanol, hoặc các axit hữu cơ như axit pyruvic, axit suxinic hoặc axit xitric có thể được sử dụng làm nguồn cacbon. Ví dụ, pepton, chất chiết xuất từ thịt, chất chiết xuất từ nấm men, chất thủy phân casein, chất chiết xuất kiềm của bột đậu nành, các alkylamin chẳng hạn như metylamin, hoặc amoniac hoặc muối của nó có thể được sử dụng làm nguồn nitơ. Ngoài ra, phosphat, cacbonat, sulfat, các muối của magie, canxi, kali, sắt, mangan, kẽm, hoặc tương tự, một axit amin cụ thể, một loại vitamin cụ thể, chất chống tạo bọt, v.v. có thể được sử dụng, nếu cần.

Việc nuôi cấy thường có thể được thực hiện, nếu cần, có khuấy hoặc lắc, ở 10 °C đến 40 °C trong 6 giờ đến 72 giờ, tốt hơn là trong 9 giờ đến 60 giờ, tốt hơn nữa là trong 12 giờ đến 48 giờ. Trong quá trình nuôi cấy, có thể thêm kháng sinh chẳng hạn như ampicillin hoặc kanamycin vào môi trường nuôi cấy, nếu cần.

Các phương pháp thu gom và tinh chế axit 4-amino-3-hydroxybenzoic từ môi trường nuôi cấy không hạn chế cụ thể. Cụ thể, việc thu gom và tinh chế có thể được thực hiện bằng cách kết hợp phương pháp nhựa trao đổi ion đã biết đến rộng rãi, phương pháp kết tủa, phương pháp kết tinh, phương pháp kết tinh lại, phương pháp cô đặc và các phương pháp khác. Ví dụ, sau khi loại bỏ các tế bào vi khuẩn bằng cách ly tâm hoặc cách tương tự, các chất ion được loại bỏ bằng nhựa trao đổi cation và anion, và sản phẩm tạo thành có thể được cô đặc để thu được axit 4-amino-3-hydroxybenzoic. Axit 4-amino-3-hydroxybenzoic được tích lũy trong môi trường nuôi cấy có thể được sử dụng trực tiếp mà không cần phân lập.

Sáng chế cũng bao gồm các chất, các phương pháp sản xuất, sử dụng, các phương pháp, v.v.. sau đây làm các phương án được lấy làm ví dụ. Tuy nhiên, sáng chế không bị giới hạn bởi các phương án này.

<1> Polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, polypeptit được chọn từ nhóm bao gồm A) đến C) sau đây:

A) polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó là leuxin;

B) polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 là phenylalanin; và

C) polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất

90% với nó, và có gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 là axit amin sau đây:

- (a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,
- (b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,
- (c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,
- (d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và
- (e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

<2> Polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, polypeptit được chọn từ nhóm bao gồm A") đến C") sau đây:

A") polypeptit đột biến có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, trong đó gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó được thay thế bằng leuxin;

B") polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 được thay thế bằng phenylalanin; và

C") polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 được thay thế bằng axit amin sau đây:

- (a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,
- (b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,

(c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,

(d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và

(e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

<3> Polypeptit đột biến theo <2>, trong đó sự thay thế gốc axit amin trong A") là sự thay thế của valin bằng leuxin, sự thay thế gốc axit amin trong B") là sự thay thế của tyrosin bằng phenylalanin, và sự thay thế gốc axit amin trong C") là sự thay thế của threonin ở vị trí 294 hoặc vị trí tương ứng với nó bằng serin, sự thay thế của valin ở vị trí 47 hoặc vị trí tương ứng với nó bằng isoleuxin, threonin, methionin hoặc glutamin, hoặc sự thay thế của histidin ở vị trí 72 hoặc vị trí tương ứng với nó bằng methionin.

<4> Phương pháp sản xuất polypeptit đột biến có hoạt tính hydroxyl hóa axit 4-aminobenzoic, bao gồm thay thế gốc axit amin, thay thế được chọn từ nhóm bao gồm A') đến C') sau đây:

A') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự thay thế gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó bằng leuxin;

B') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự thay thế gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 bằng phenylalanin; và

C') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự thay thế gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 bằng axit amin sau đây:

(a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,

(b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,

(c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,

(d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và

(e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

<5> Phương pháp theo <4>, trong đó sự thay thế gốc axit amin trong A') là sự thay thế của valin bằng leuxin, sự thay thế của gốc axit amin trong B') là sự thay thế của tyrosin bằng phenylalanin, và sự thay thế gốc axit amin trong C') là sự thay thế của threonin ở vị trí 294 hoặc vị trí tương ứng với nó bằng serin, sự thay thế của valin ở vị trí 47 hoặc vị trí tương ứng với nó bằng isoleuxin, threonin, methionin hoặc glutamin, hoặc sự thay thế của histidin ở vị trí 72 hoặc vị trí tương ứng với nó bằng methionin.

<6> Phương pháp cải thiện hoạt tính hydroxyl hóa axit 4-aminobenzoic, bao gồm thay thế gốc axit amin, thay thế được chọn từ nhóm bao gồm A') đến C') sau đây:

A') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự thay thế gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó bằng leuxin;

B') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự thay thế gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 bằng phenylalanin; và

C') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự thay thế gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 bằng axit amin sau đây:

(a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,

(b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,

(c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,

(d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và

(e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

<7> Phương pháp theo <6>, trong đó sự thay thế của gốc axit amin trong A') là sự thay thế của valin bằng leuxin, sự thay thế gốc axit amin trong B') là sự thay thế của tyrosin bằng phenylalanin, và sự thay thế gốc axit amin trong C') là sự thay thế của threonin ở vị trí 294 hoặc vị trí tương ứng với nó bằng serin, sự thay thế của valin ở vị trí 47 hoặc vị trí tương ứng với nó bằng isoleuxin, threonin, methionin hoặc glutamin, hoặc sự thay thế của histidin ở vị trí 72 hoặc vị trí tương ứng với nó bằng methionin.

<8> Phương pháp cải thiện năng suất của axit 4-amino-3-hydrozybenzoic, bao gồm thay thế gốc axit amin, thay thế được chọn từ nhóm bao gồm A') đến C') sau đây:

A') trong trường hợp sản xuất axit 4-amino-3-hydroxybenzoic bằng cách sử dụng polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 47% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự thay thế gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó bằng leuxin;

B') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 51% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự thay thế gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 bằng phenylalanin; và

C') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, sự thay thế gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385

của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 bằng axit amin sau đây:

- (a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,
- (b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,
- (c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,
- (d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và
- (e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

<9> Phương pháp theo <8>, trong đó sự thay thế gốc axit amin trong A') là sự thay thế của valin bằng leuxin, sự thay thế gốc axit amin trong B') là sự thay thế của tyrosin bằng phenylalanin, và sự thay thế gốc axit amin trong C') là sự thay thế của threonin ở vị trí 294 hoặc vị trí tương ứng với nó bằng serin, sự thay thế của valin ở vị trí 47 hoặc vị trí tương ứng với nó bằng isoleuxin, threonin, methionin hoặc glutamin, hoặc sự thay thế của histidin ở vị trí 72 hoặc vị trí tương ứng với nó bằng metionin.

<10> Polynucleotit mã hóa polypeptit theo bất kỳ từ <1> đến <3>.

<11> Vector hoặc mảnh ADN bao gồm polynucleotit theo <10>.

<12> Tế bào được biến nạp bao gồm vector hoặc mảnh ADN theo <11>.

<13> Tế bào được biến nạp theo <12>, trong đó tế bào được biến nạp có nguồn gốc từ *E. coli* hoặc vi khuẩn thuộc giống *Corynebacterium*.

<14> Tế bào được biến nạp theo <12> hoặc <13>, trong đó tế bào được biến nạp là vi sinh vật có khả năng cung cấp axit 4-aminobenzoic.

<15> Tế bào được biến nạp theo <12> hoặc <13>, trong đó tế bào được biến nạp đã cải thiện khả năng cung cấp axit 4-aminobenzoic.

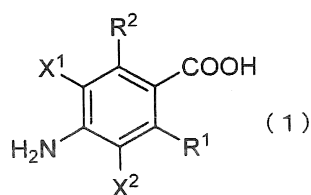
<16> Phương pháp sản xuất axit 4-amino-3-hydroxybenzoic, bao gồm bước nuôi cấy tế bào được biến nạp theo bất kỳ từ <12> đến <15>.

<17> Phương pháp theo <16>, trong đó việc nuôi cấy được thực hiện trong môi trường nuôi cấy chứa sacarit làm nguồn cacbon.

<18> Phương pháp theo <16> hoặc <17>, bao gồm thêm bước thu gom axit 4-amino-3-hydroxybenzoic từ môi trường nuôi cấy.

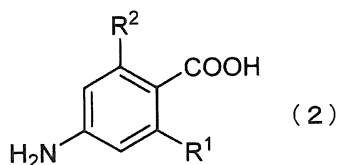
<19> Phương pháp theo bất kỳ từ <16> đến <18>, trong đó việc nuôi cấy được thực hiện với sự có mặt của axit 4-aminobenzoic.

<20> Phương pháp theo bất kỳ từ <16> đến <19>, trong đó axit 4-amino-3-hydroxybenzoic là các dẫn xuất axit 4-amino-3-hydroxybenzoic có công thức (1) sau đây:



trong đó R^1 là nguyên tử hydro, nhóm hydroxy, nhóm metoxy, nhóm amin, nguyên tử flo, nguyên tử clo, nguyên tử brom, nguyên tử iốt, nhóm cacboxy, nhóm methyl, hoặc nhóm etyl, R^2 là nguyên tử hydro, nhóm hydroxy, nhóm metoxy, nhóm amin, nguyên tử flo, nguyên tử clo, nguyên tử brom, nguyên tử iốt, nhóm cacboxy, nhóm methyl, hoặc nhóm etyl, và X^1 và X^2 mỗi loại là nguyên tử hydro hoặc nhóm hydroxy, ít nhất một trong số đó là nhóm hydroxy.

<21> Phương pháp theo <19> hoặc <20>, trong đó axit 4-aminobenzoic là các dẫn xuất axit 4-aminobenzoic có công thức (2) sau đây:



trong đó R^1 là nguyên tử hydro, nhóm hydroxy, nhóm metoxy, nhóm amin, nguyên tử flo, nguyên tử clo, nguyên tử brom, nguyên tử iốt, nhóm cacboxy, nhóm methyl, hoặc nhóm etyl, và R^2 là nguyên tử hydro, nhóm hydroxy, nhóm metoxy, nhóm amin, nguyên tử flo, nguyên tử clo, nguyên tử brom, nguyên tử iốt, nhóm cacboxy, nhóm methyl, hoặc nhóm etyl.

Ví dụ thực hiện sáng chế

Sau đây, sáng chế sẽ được mô tả chi tiết hơn có tham chiếu đến các ví dụ. Tuy nhiên, sáng chế này không bị giới hạn.

Ví dụ A1: Sản xuất axit 4-amino-3-hydroxybenzoic

Trong ví dụ sau đây, PCR được thực hiện bằng PrimeSTAR Max Premix (Takara Bio Inc.).

(1) Chuẩn bị plasmit chứa gen mã hóa enzym kiểu đại

(a) Chuẩn bị plasmit pECsf_gapS_pabABC

Mảnh ADN chứa các gen mã hóa 4-amino-4-deoxychorismat synthaza và 4-amino-4-deoxychorismat lyaza được khuếch đại bằng PCR sử dụng bộ gen được tách chiết từ chủng *Corynebacterium glutamicum* ATCC13032 bằng phương pháp thông thường làm khuôn và sử dụng các đoạn mồi GN14_127 (SEQ ID NO: 13, TATTAATTAAATGCGCGTTTAAATTATTGATAATTATGATTC) và GN14_133 (SEQ ID NO: 14, TTGCGGCCGCTTGTTTAAACCTCCTTACAGAAAAATGGTTGGGCG). Đoạn này được chèn vào giữa vị trí PacI và vị trí NotI của plasmit pECsf_gapS (xem Đơn sáng chế Nhật Bản số 2015-25491) để thu được plasmit pECsf_gapS_pabABC.

(b) Chuẩn bị plasmit pECsf_gapS_pabABC_HFM122

Mảnh ADN cho một vectơ được tổng hợp bằng PCR sử dụng plasmit pECsf_gapS_pabABC thu được ở trên làm khuôn và sử dụng các đoạn mồi pabABCcory vec R (SEQ ID NO: 15, AAATTTAAACCTCCTTTACAGAAAAATGGTTGG) và pabABCcory vec F (SEQ ID NO: 16, GGAGGTTTAAACAAGCGGCCGCGATATC). Sau đó, plasmit chứa gen (SEQ ID NO: 1) mã hóa polypeptit HFM122 có hoạt tính hydroxyl hóa axit 4-aminobenzoic được chuẩn bị bằng cách tổng hợp gen nhân tạo, và mảnh ADN cho đoạn chèn được tổng hợp bằng PCR sử dụng plasmit này làm khuôn và sử dụng các đoạn mồi pECsfD HFM122 F (SEQ ID NO: 17, AGGAGGTTTAAATTTATGCGCACTCAGGTGGCTAT) và pECsfD HFM122 R (SEQ ID NO: 18, CTTGTTTAAACCTCCTTATACGAGTGGCAGTCCTA). Các sản

phẩm PCR này được xử lý bằng DpnI (Takara Bio Inc.). Sau đó, các mảnh ADN tương ứng được tinh chế bằng cách sử dụng NucleoSpin Gel và PCR Clean-up (Takara Bio Inc.) và được gắn bằng cách sử dụng Kit tách dòng In-Fusion HD (Takara Bio Inc.) để xây dựng plasmit pECsf_gapS_pabABC_HFM122. Chủng *E. coli* DH5 α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch plasmit thu được. Dung dịch tế bào được cấy trên môi trường thạch LBKm (1 % Bacto Trypton, 0,5 % yeast extract, 1 % NaCl, 50 μ g/mL kanamycin sulfat, 1,5 % thạch) và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được phải chịu phản ứng PCR sử dụng Sapphire Amp (Takara Bio Inc.) và các đoạn mồi pabABC+pobA cho CPCR F (SEQ ID NO: 19, GCTATCAAAACATTCGGCACATTGGTTTTCC) và pabABC+pobA cho CPCR R (SEQ ID NO: 20, GGAAGATGCGTGATCTGATCCTTCAACTC) để chọn một thể biến nạp được xác nhận là chứa mảnh ADN quan tâm. Thể biến nạp thu được được cấy vào 2 mL môi trường lỏng LBKm (1 % Bacto Trypton, 0,5 % chiết xuất nấm men, 1 % NaCl, 50 μ g/mL kanamycin sulfat) và nuôi cấy qua đêm ở 37 °C. Plasmit được tinh chế từ dung dịch nuôi cấy này bằng cách sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

(c) Chuẩn bị plasmit pECsf_gapS_pabABC_tuD_HFM122

Mảnh ADN cho vectơ được tổng hợp bằng PCR sử dụng plasmit pECsf_gapS_pabABC_HFM122 thu được ở trên làm khuôn và sử dụng các đoạn mồi pabC last R (SEQ ID NO: 21, TTACAGAAAAATGGTTGGGCGCAA) và HFM122 F (SEQ ID NO: 22, ATGCGCACTCAGGTGGCTATCG). Sau đó, mảnh ADN (SEQ ID NO: 23, TACGTACCTGCAGGTAGCGTGTCTAGTAGGCGCGTAGGGTAAGTGGGGTAG CGGCTTGTTAGATATCTTGAAATCGGCTTTCAACAGCATTGATTTTCGATGT ATTTAGCTGGCCGTTACCCTGCGAATGTCCACAGGGTAGCTGGTAGTTTGA AAATCAACGCCGTTGCCCTTAGGATTCAGTAACTGGCACATTTTGTAAATGC GCTAGATCTGTGTGCTCAGTCTTCCAGGCTGCTTATCACAGTGAAAGCAAA ACCAATTTCGTGGCTGCGAAAGTCGTAGCCACCACGAAGTCCAAAGGAGGA TCTAAATTATGAATAATATAAAAGGAGGAATTAATTAA) chứa trình tự khởi

động gen *tuf* (cg0587) (sau đây, được gọi là trình tự khởi động *tuf*) được mang bởi chủng *Corynebacterium glutamicum* ATCC13032 được chuẩn bị bằng cách tổng hợp gen nhân tạo, và mảnh ADN cho đoạn chèn được tổng hợp bằng PCR sử dụng đoạn này làm khuôn và sử dụng các đoạn mồi *pabC*-*Ptu* F (SEQ ID NO: 24, ACCATTTTTCTGTAATACGTACCTGCAGGTAGCGTG) và *Ptu*-HFM122 R (SEQ ID NO: 25, CACCTGAGTGCGCATTTAATTAATTCCTCCTTTTA). Các sản phẩm PCR này được xử lý bằng *DpnI* (Takara Bio Inc.). Sau đó, các mảnh ADN tương ứng được tinh chế bằng cách sử dụng NucleoSpin Gel và PCR Clean-up (Takara Bio Inc.) và được gắn bằng cách sử dụng Kit tách dòng In-Fusion HD (Takara Bio Inc.) để xây dựng plasmit *pECsf_gapS_pabABC_tuD_HFM122*. Chủng *E. coli* DH5 α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch plasmit thu được. Dung dịch tế bào được cấy trên môi trường thạch LBKm và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được đã trải qua phản ứng PCR sử dụng Sapphire Amp (Takara Bio Inc.) và các đoạn mồi *Ptu* seq 1 (SEQ ID NO: 26, GCTTGTTAGATATCTTGAAATCGGCTTTC) và *pabABC*+*pobA* cho CPCR R (SEQ ID NO: 20, GGAAGATGCGTGATCTGATCCTTCAACTC) để chọn thể biến nạp được xác nhận chứa mảnh ADN quan tâm. Thể biến nạp thu được được cấy vào 2 mL môi trường lỏng LBKm và được nuôi cấy qua đêm ở 37 °C. Plasmit được tinh chế từ dung dịch nuôi cấy này bằng cách sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

Trong plasmit được xây dựng, các gen mã hóa 4-amino-4-deoxychorismat synthaza và 4-amino-4-deoxychorismat lyaza được liên kết dưới sự kiểm soát của trình tự khởi động *gap*, và gen mã hóa HFM122 kiểu đại được liên kết thêm dưới sự kiểm soát của trình tự khởi động *tuf*.

(d) Chuẩn bị các plasmit khác

Mảnh ADN cho vector được tổng hợp bằng PCR sử dụng plasmit *pECsf_gapS_pabABC_tuD_HFM122* thu được ở trên làm khuôn và sử dụng các đoạn mồi *pGapABA_tu* vec F (SEQ ID NO: 27, GGAGGTTTAAACAAGCGG) và *pGapABA_tu* vec R (SEQ ID NO: 28, AATTTAGATCCTCCTTTGGACTTCGTG).

Sau đó, plasmit chứa gen (SEQ ID NO: 3, 5, 7, 9, hoặc 11) mã hóa mỗi polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic đã được chuẩn bị bằng cách tổng hợp gen nhân tạo, và mảnh ADN cho đoạn chèn được tổng hợp bằng PCR sử dụng plasmit này làm khuôn và sử dụng các đoạn mồi được thể hiện trong cột "Đoạn mồi" của Bảng A1. Các sản phẩm PCR này được xử lý bằng DpnI (Takara Bio Inc.). Sau đó, các mảnh ADN tương ứng được tinh chế bằng cách sử dụng NucleoSpin Gel và PCR Clean-up (Takara Bio Inc.) và được gắn bằng cách sử dụng Kit tách dòng In-Fusion HD (Takara Bio Inc.) để xây dựng plasmit được thể hiện trong cột "Plasmit" của Bảng A1. Chúng *E. coli* DH5 α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch plasmit thu được. Dung dịch tế bào được cấy trên môi trường thạch LBKm và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được đã trải qua phản ứng PCR sử dụng Sapphire Amp (Takara Bio Inc.) và các đoạn mồi Ptu seq 1 (SEQ ID NO: 26, GCTTGTTAGATATCTTGAAATCGGCTTTC) và pabABC+pobA cho CPCR R (SEQ ID NO: 20, GGAAGATGCGTGATCTGATCCTTCAACTC) để chọn thể biến nạp được xác nhận chứa mảnh ADN quan tâm. Thể biến nạp thu được được cấy vào 2 mL môi trường lỏng LBKm và được nuôi cấy qua đêm ở 37 °C. Plasmit đã được tinh chế từ dung dịch nuôi cấy này bằng cách sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

Trong các plasmit được xây dựng, các gen mã hóa 4-amino-4-deoxychorismat synthaza và 4-amino-4-deoxychorismat lyaza được liên kết dưới sự kiểm soát của trình tự khởi động gap, và gen mã hóa hydroxylaza kiểu dại được liên kết thêm dưới sự kiểm soát của trình tự khởi động tu.

[Bảng A1]

Plasmit	Đoạn mồi	Trình tự (5'→3')	SEQ ID NO
pECsf_gapS_pabABC_tu_HFM77	HFM77 ins F	AGGAGGATCTAAATTATGCGTAC TCAGGTGGGAATC	29
	HFM77 ins R	CTTGTTTAAACCTCCTTAAGCAA GTGGCATGC	30
pECsf_gapS_pabABC_tu_HFM339	HFM339 ins F	AGGAGGATCTAAATTATGCGCAC TCAGGTGGCAATC	31
	HFM339 ins R	CTTGTTTAAACCTCCTTAGTATGG CAGGCCTACG	32
pECsf_gapS_pabABC_tu_HFM388	HFM388 ins F	AGGAGGATCTAAATTATGCGCAC CCAAGTGGTCATC	33
	HFM388 ins R	CTTGTTTAAACCTCCTTAGAACG GCAGACCCACGTAG	34
pECsf_gapS_pabABC_tu_HFM737	HFM737 ins F	AGGAGGATCTAAATTATGCGCAC TCAGGTTGGTATC	35
	HFM737 ins R	CTTGTTTAAACCTCCTTAGTGGCT CAGTCCAACCATTC	36
pECsf_gapS_pabABC_tu_HFMss0-1	HFMss0-1 ins F	AGGAGGATCTAAATTATGCGTAC CCAAGTGGCCATCATTG	37
	HFMss0-1 ins R	CTTGTTTAAACCTCCTTAGAAGCC AATCGGAAGGCC	38

(2) Chuẩn bị plasmit chứa gen mã hóa enzym đột biến

Việc chuẩn bị plasmit chứa gen mã hóa đột biến enzym sẽ được đưa ra dưới đây bằng cách lấy làm ví dụ, chuẩn bị plasmit chứa gen mã hóa enzym đột biến trong đó valin ở vị trí 47 của HFM77 được thay thế bằng leuxin.

Plasmit pECsf_gapS_pabABC_tu_HFM77_V47L được xây dựng bằng PCR sử dụng plasmit pECsf_gapS_pabABC_tu_HFM77 làm khuôn và sử dụng các đoạn môi bổ sung HFM77 V47L F (SEQ ID NO: 39, GCCGGGCTCCTGGAACAGTCTACGGTT) và HFM77 V47L R (SEQ ID NO: 40, TTCCAGGAGCCCGGCGCGGATGGTCTG). Sản phẩm PCR được xử lý bằng DpnI (Takara Bio Inc.). Chúng E. coli DH5 α của ECOS Competent (Nippon Gene Co., Ltd.) đã được biến nạp bằng dung dịch do đó đã được xử lý. Dung dịch tế bào được cấy trên môi trường thạch LBKm và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được được chọn làm thể biến nạp. Thể biến nạp được cấy vào 2 mL môi trường lỏng LBKm và nuôi cấy qua đêm ở 37 °C. Plasmit đã được tinh chế từ dung dịch nuôi cấy này bằng cách sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

Theo cách tương tự, plasmit chứa gen mã hóa mỗi đột biến enzym thu được bằng PCR sử dụng plasmit được thể hiện trong cột "Khuôn" của Bảng A2 thay cho plasmit pECsf_gapS_pabABC_tu_HFM77 và sử dụng các đoạn môi được thể hiện trong cột "Đoạn môi" của Bảng A2 thay cho các đoạn môi HFM77 V47L F và HFM77 V47L R.

[Bảng A2]

Khuôn	Đoạn mồi	Trình tự (5'→3')	SEQ ID NO
pECsf_gapS_pabABC_tu D_HFM122	HFM122 V47L F	GCTGGTCTCCTGGAACGTATC ACGGTG	41
	HFM122 V47L R	TTCCAGGAGACCAGCCCGAA CTCGGCC	42
pECsf_gapS_pabABC_tu _HFM339	HFM339 V47L F	GCAGGCCTCCTGGAGCAGGG CATGGTT	43
	HFM339 V47L R	CTCCAGGAGGCCTGCACGGA TGC GGGA	44
pECsf_gapS_pabABC_tu _HFM388	HFM388 V47L F	GCTGGACTCTTGGAACAGGG CACCGTT	45
	HFM388 V47L R	TTCCAAGAGTCCAGCGCGAA CTCGCCC	46
pECsf_gapS_pabABC_tu _HFM737	HFM737 V47L F	GCGGGTCTCCTGGAACAGGG CACCATG	47
	HFM737 V47L R	TTCCAGGAGACCCGCCCCGAA TCGTGGA	48
pECsf_gapS_pabABC_tu _HFMss0-1	HFMss0-1 V47L F	GCCGGTCTCCTGGAGCACTCC ACGGTG	49
	HFMss0-1 V47L R	CTCCAGGAGACCGGCACGGA TGC GTGA	50

(3) Đưa plasmit vào tế bào chủ

Chủng *Corynebacterium glutamicum* DRHG145 (xem Đơn sáng chế Nhật Bản số 2014-523757) được biến nạp với mỗi plasmit thu được ở trên bằng công nghệ tạo lỗ xung điện (Bio-Rad Laboratories, Inc.). Dung dịch tế bào được biến nạp thu được được cấy trên môi trường thạch LBKm và sau đó để yên ở 30 °C trong 2 ngày. Các khuẩn lạc thu được này được sử dụng làm thể biến nạp.

(4) Nuôi cấy thể biến nạp

Mỗi thể biến nạp thu được ở trên được cấy vào 1 mL môi trường CGYE (chứa 50 µg/mL kanamycin sulfat) được thể hiện trong Bảng A3, và được nuôi cấy qua đêm ở 30 °C. 100 µL dung dịch nuôi cấy thu được được cấy vào 10 mL môi trường CGXII (chứa 50 µg/mL kanamycin sulfat) được thể hiện trong Bảng A4, và được nuôi cấy ở 30 °C trong khoảng 48 giờ. Sau đó, các tế bào vi khuẩn được loại bỏ bằng cách ly tâm để thu được dịch nổi nuôi cấy. Nồng độ của axit 4-amino-3-hydroxybenzoic trong dịch nổi nuôi cấy thu được được định lượng theo phương pháp của Ví dụ tham chiếu 1. Tỷ lệ cải thiện khả năng sản xuất axit 4-amino-3-hydroxybenzoic được tính theo phương trình đưa ra dưới đây. Trong phương trình, "WT" là "thể biến nạp chứa plasmit chứa gen mã hóa enzym kiểu dại", và "MT" là "thể biến nạp chứa plasmit chứa gen mã hóa enzym đột biến và được chuẩn bị từ plasmit chứa gen mã hóa enzym kiểu dại đã được đưa vào".

(Phương trình 1)

Tỷ lệ cải thiện khả năng sản xuất = Khả năng của MT sản xuất axit 4-amino-3-hydroxybenzoic / Khả năng của WT sản xuất axit 4-amino-3-hydroxybenzoic

[Bảng A3]

Thành phần môi trường CGYE (mỗi L)

Glucosa	50 g
$(\text{NH}_4)_2\text{SO}_4$	20 g
<u>Ure</u>	5 g
KH_2PO_4	1 g
K_2HPO_4	1 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0,25 g
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	10 mg
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	10 mg
$\text{MnSO}_4 \cdot 5\text{H}_2\text{O}$	10 mg
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	1 mg
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	0,2 mg
$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$	0,02 mg
Biotin (pH7)	0,2 mg
Dịch chiết nấm men	1 g

[Bảng A4]

Thành phần môi trường CGXII (mỗi L)

Glucoza	50 g
$(\text{NH}_4)_2\text{SO}_4$	20 g
<u>Ure</u>	5 g
KH_2PO_4	1 g
K_2HPO_4	1 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0,25 g
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	10 mg
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	10 mg
$\text{MnSO}_4 \cdot 5\text{H}_2\text{O}$	10 mg
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	1 mg
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	0,2 mg
$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$	0,02 mg
Biotin (pH7)	0,2 mg
<u>Trypton</u>	10 g

(5) Kết quả

Như được thể hiện trong Bảng A5, chủng vi khuẩn chứa mỗi enzym đột biến có khả năng sản xuất axit 4-amino-3-hydroxybenzoic được cải thiện nhiều hơn so với chủng vi khuẩn chứa enzym kiểu dại.

[Bảng A5]

Hydroxylaza	Khả năng sản xuất axit 4-amino-3-hydroxybenzoic (g/L)	Tỷ lệ cải thiện khả năng sản xuất
HFM77 wt	0,073	1,00
HFM77 V47L	0,134	1,83
HFM122 wt	0,134	1,00
HFM122 V47L	0,372	2,77
HFM339 wt	0,016	1,00
HFM339 V47L	0,071	4,47
HFM388 wt	0,033	1,00
HFM388 V47L	0,100	3,02
HFM737 wt	0,112	1,00
HFM737 V47L	0,164	1,46
HFMss0-1 wt	0,112	1,00
HFMss0-1 V47L	0,140	1,25

Ví dụ B1: Sản xuất axit 4-amino-3-hydroxybenzoic

Trong ví dụ sau, PCR được thực hiện sử dụng PrimeSTAR Max Premix (Takara Bio Inc.).

(1) Chuẩn bị plasmit chứa gen mã hóa enzym kiểu đại

(a) Chuẩn bị plasmit pECsf_gapS_pabABC

Mảnh ADN chứa các gen mã hóa 4-amino-4-deoxychorismat synthaza và 4-amino-4-deoxychorismat lyaza được khuếch đại bằng PCR sử dụng bộ gen chiết xuất từ chủng *Corynebacterium glutamicum* ATCC13032 bằng phương pháp thông thường làm khuôn và sử dụng các đoạn mồi GN14_127 (SEQ ID NO: 13, TATTAATTAAATGCGCGTTTAAATTATTGATAATTATGATTC) và GN14_133 (SEQ ID NO: 14, TTGCGGCCGCTTGTTTAAACCTCCTTACAGAAAAATGGTTGGGCG). Đoạn này được chèn vào giữa vị trí PacI và vị trí NotI của plasmit pECsf_gapS (xem Đơn sáng chế Nhật Bản số 2015-25491) để thu được plasmit pECsf_gapS_pabABC.

(b) Chuẩn bị plasmit pECsf_gapS_pabABC_HFM122

Một mảnh ADN cho vector được tổng hợp bằng PCR sử dụng plasmit pECsf_gapS_pabABC thu được ở trên làm khuôn và sử dụng các đoạn mồi pabABCcory vec R (SEQ ID NO: 15, AAATTTAAACCTCCTTTACAGAAAAATGGTTGG) và pabABCcory vec F (SEQ ID NO: 16, GGAGGTTTAAACAAGCGGCCGCGATATC). Sau đó, plasmit chứa gen (SEQ ID NO: 1) mã hóa polypeptit HFM122 có hoạt tính hydroxyl hóa axit 4-aminobenzoic được chuẩn bị bằng cách tổng hợp gen nhân tạo, và một mảnh ADN cho đoạn chèn được tổng hợp bằng PCR sử dụng plasmit này làm khuôn và sử dụng các đoạn mồi pECsfD HFM122 F (SEQ ID NO: 17, AGGAGGTTTAAATTTATGCGCACTCAGGTGGCTAT) và pECsfD HFM122 R (SEQ ID NO: 18, CTTGTTTAAACCTCCTTATACGAGTGGCAGTCCTA). Các sản phẩm PCR này được xử lý bằng DpnI (Takara Bio Inc.). Sau đó, các mảnh ADN tương ứng được tinh chế bằng cách sử dụng NucleoSpin Gel và PCR Clean-up (Takara Bio Inc.) và được gắn bằng cách sử dụng Kit tách dòng In-Fusion HD (Takara Bio Inc.) để

xây dựng plasmit pECsf_gapS_pabABC_HFM122. Chúng *E. coli* DH5 α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch plasmit thu được. Dung dịch tế bào được cấy trên môi trường thạch LBKm (1 % Bacto Trypton, 0,5 % chiết xuất nấm men, 1 % NaCl, 50 μ g/mL kanamycin sulfat, 1,5 % aga) và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được phải chịu phản ứng PCR sử dụng Sapphire Amp (Takara Bio Inc.) và các đoạn mồi pabABC+pobA for CPCR F (SEQ ID NO: 19, GCTATCAAAACATTCGGCACATTGGTTTTCC) và pabABC+pobA cho CPCR R (SEQ ID NO: 20, GGAAGATGCGTGATCTGATCCTTCAACTC) để chọn thể biến nạp được xác nhận là chứa mảnh ADN quan tâm. Thể biến nạp thu được được cấy vào 2 mL môi trường lỏng LBKm (1 % Bacto Trypton, 0,5 % chiết xuất nấm men, 1 % NaCl, 50 μ g/mL kanamycin sulfat) và được nuôi cấy qua đêm ở 37 °C. Plasmit được tinh chế từ dung dịch nuôi cấy này sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

(c) Chuẩn bị plasmit pECsf_gapS_pabABC_tuD_HFM122

Mảnh ADN cho vector được tổng hợp bằng PCR sử dụng plasmit pECsf_gapS_pabABC_HFM122 thu được ở trên làm khuôn và sử dụng các đoạn mồi pabC last R (SEQ ID NO: 21, TTACAGAAAAATGGTTGGGCGCAA) và HFM122 F (SEQ ID NO: 22, ATGCGCACTCAGGTGGCTATCG). Sau đó, một mảnh ADN (SEQ ID NO: 23, TACGTACCTGCAGGTAGCGTGTCTAGGCGCGTAGGGTAAGTGGGGTAG CGGCTTGTTAGATATCTTGAAATCGGCTTCAACAGCATTGATTTTCGATGT ATTTAGCTGGCCGTTACCCTGCGAATGTCCACAGGGTAGCTGGTAGTTTGA AAATCAACGCCGTTGCCCTTAGGATTCAGTAACTGGCACATTTTGTAATGC GCTAGATCTGTGTGCTCAGTCTTCCAGGCTGCTTATCACAGTGAAAGCAAA ACCAATTCGTGGCTGCGAAAGTCGTAGCCACCACGAAGTCCAAAGGAGGA TCTAAATTATGAATAATATAAAAGGAGGAATTAATTAA) chứa trình tự khởi động gen tuf (cg0587) (sau đây, được gọi là trình tự khởi động tu) được mang bởi chủng *Corynebacterium glutamicum* ATCC13032 được chuẩn bị bằng cách tổng hợp gen nhân tạo, và mảnh ADN cho đoạn chèn được tổng hợp bằng PCR sử dụng đoạn

này làm khuôn và sử dụng các đoạn môi pabC-Ptu F (SEQ ID NO: 24, ACCATTTTCTGTAATACGTACCTGCAGGTAGCGTG) và Ptu-HFM122 R (SEQ ID NO: 25, CACCTGAGTGCGCATTTAATTAATTCCTCCTTTTA). Các sản phẩm PCR này được xử lý bằng DpnI (Takara Bio Inc.). Sau đó, các mảnh ADN tương ứng được tinh chế bằng cách sử dụng NucleoSpin Gel và PCR Clean-up (Takara Bio Inc.) và được gắn vào bằng cách sử dụng Kit tách dòng In-Fusion HD (Takara Bio Inc.) để xây dựng plasmid pECsf_gapS_pabABC_tuD_HFM122. Chúng *E. coli* DH5 α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch plasmid thu được. Dung dịch tế bào được cấy trên môi trường thạch LBKm và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được đã trải qua phản ứng PCR sử dụng Sapphire Amp (Takara Bio Inc.) và các đoạn môi Ptu seq 1 (SEQ ID NO: 26, GCTTGTTAGATATCTTGAAATCGGCTTTC) và pabABC+pobA cho CPCR R (SEQ ID NO: 20, GGAAGATGCGTGATCTGATCCTTCAACTC) để chọn thể biến nạp được xác nhận chứa mảnh ADN quan tâm. Thể biến nạp thu được được cấy vào 2 mL môi trường lỏng LBKm và được nuôi cấy qua đêm ở 37 °C. Plasmid được tinh chế từ dung dịch nuôi cấy này sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

Trong plasmid được xây dựng, các gen mã hóa 4-amino-4-deoxychorismat synthaza và 4-amino-4-deoxychorismat lyaza được liên kết dưới sự kiểm soát của trình tự khởi động gap, và gen mã hóa HFM122 kiểu dại được liên kết thêm dưới sự kiểm soát của trình tự khởi động tu.

(d) Chuẩn bị các plasmid khác

Mảnh ADN cho vector được tổng hợp bằng PCR sử dụng plasmid pECsf_gapS_pabABC_tuD_HFM122 thu được ở trên làm khuôn và sử dụng các đoạn môi pGapABA_tu vec F (SEQ ID NO: 27, GGAGGTTTAAACAAGCGG) và pGapABA_tu vec R (SEQ ID NO: 28, AATTTAGATCCTCCTTTGGACTTCGTG). Sau đó, plasmid chứa gen (SEQ ID NO: 3, 5, hoặc 7) mã hóa mỗi polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic được chuẩn bị bằng cách tổng hợp gen nhân tạo, và một mảnh ADN cho đoạn chèn được tổng hợp bằng PCR sử dụng plasmid này làm khuôn và sử dụng các đoạn môi được thể hiện trong cột "Đoạn môi" của Bảng B1.

Các sản phẩm PCR này được xử lý bằng DpnI (Takara Bio Inc.). Sau đó, các mảnh ADN tương ứng được tinh chế bằng cách sử dụng NucleoSpin Gel và PCR Clean-up (Takara Bio Inc.) và được gắn bằng cách sử dụng Kit tách dòng In-Fusion HD (Takara Bio Inc.) để xây dựng plasmid được thể hiện trong cột "Plasmid" của Bảng B1. Chủng *E. coli* DH5 α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch plasmid thu được. Dung dịch tế bào được cấy trên môi trường thạch LBKm và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được đã trải qua phản ứng PCR sử dụng Sapphire Amp (Takara Bio Inc.) và các đoạn mồi Ptu seq 1 (SEQ ID NO: 26, GCTTGTTAGATATCTTGAAATCGGCTTTC) và pabABC+pobA cho CPCR R (SEQ ID NO: 20, GGAAGATGCGTGATCTGATCCTTCAACTC) để chọn thể biến nạp được xác nhận chứa mảnh ADN quan tâm. Thể biến nạp thu được được cấy vào 2 mL môi trường lỏng LBKm và được nuôi cấy qua đêm ở 37 °C. Plasmid được tinh chế từ dung dịch nuôi cấy này bằng cách sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

Trong các plasmid được xây dựng, các gen mã hóa 4-amino-4-deoxychorismat synthaza và 4-amino-4-deoxychorismat lyaza được liên kết dưới sự kiểm soát của trình tự khởi động gap, và gen mã hóa hydroxylaza kiểu đại được liên kết thêm dưới sự kiểm soát của trình tự khởi động tu.

[Bảng B1]

Plasmit	Đoạn môi	Trình tự (5'→3')	SEQ ID NO
pECsf_gapS_pabA BC_tu_HFM77	HFM77 ins F	AGGAGGATCTAAATTATGCGTA CTCAGGTGGGAATC	29
	HFM77 ins R	CTTGTTTAAACCTCCTTAAGCA AGTGGCATGC	30
pECsf_gapS_pabA BC_tu_HFM339	HFM339 ins F	AGGAGGATCTAAATTATGCGCA CTCAGGTGGCAATC	31
	HFM339 ins R	CTTGTTTAAACCTCCTTAGTATG GCAGGCCTACG	32
pECsf_gapS_pabA BC_tu_HFM388	HFM388 ins F	AGGAGGATCTAAATTATGCGCA CCCAAGTGGTCATC	33
	HFM388 ins R	CTTGTTTAAACCTCCTTAGAAC GGCAGACCCACGTAG	34

(2) Chuẩn bị plasmit chứa gen mã hóa enzym đột biến

Việc chuẩn bị plasmit chứa gen mã hóa đột biến sẽ được đưa ra dưới đây bằng cách lấy làm ví dụ, chuẩn bị plasmit chứa gen mã hóa đột biến enzym trong đó tyrosin ở vị trí 201 của HFM77 được thay thế bằng phenylalanin.

Plasmit pECsf_gapS_pabABC_tu_HFM77_Y201F đã được xây dựng bằng PCR sử dụng plasmit pECsf_gapS_pabABC_tu_HFM77 làm khuôn và sử dụng các đoạn môi bổ sung HFM77 Y201F F (SEQ ID NO: 51, CTCATCTTCGCACATCACGACCGCGGA) và HFM77 Y201F R (SEQ ID NO: 52, ATGTGCGAAGATGAGCTCTTCGGATGA). Sản phẩm PCR được xử lý bằng DpnI (Takara Bio Inc.). Chúng E. coli DH5α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch do đó đã được xử lý. Dung dịch tế bào được cấy trên môi trường thạch LBKm và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được được chọn làm thể biến nạp. Thể biến nạp được cấy vào 2 mL môi trường lỏng

LBKm và được nuôi cấy qua đêm ở 37 °C. Plasmit được tinh chế từ dung dịch nuôi cấy này bằng cách sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

Theo cách tương tự, plasmit chứa gen mã hóa mỗi đột biến enzym thu được bằng PCR sử dụng plasmit được thể hiện trong cột "Khuôn" của Bảng B2 thay cho plasmit pECsf_gapS_pabABC_tu_HFM77 và sử dụng các đoạn mồi được thể hiện trong cột "Đoạn mồi" của Bảng B2 thay cho các đoạn mồi HFM77 Y201F F và HFM77 Y201F R.

[Bảng B2]

Khuôn	Đoạn mồi	Trình tự (5'→3')	SEQ ID NO
pECsf_gapS_pabABC_tu_HFM77	HFM77 Y222F F	CGCATGTTCTTCCAGTGCGC ACCTACC	53
	HFM77 Y222F R	CTGGAAGAACATGCGCTGGA TATTCGG	54
pECsf_gapS_pabABC_tu_D_HFM122	HFM122 Y201F F	TTGATCTTCTCGAACCATGA TCGCGGT	55
	HFM122 Y201F R	GTTCGAGAAGATCAACTCGT GGTCACA	56
pECsf_gapS_pabABC_tu_D_HFM122	HFM122 Y222F F	CGCTATTTTCGTGCAGTGCTC ACTCGAC	57
	HFM122 Y222F R	CTGCACGAAATAGCGGGAG CGTGTCGG	58
pECsf_gapS_pabABC_tu_HFM339	HFM339 Y201F F	CTGATCTTCGTCAACCACGA CCGAGGC	59
	HFM339 Y201F R	GTTGACGAAGATCAGTTCTG GGCTGAC	60
pECsf_gapS_pabABC_tu_HFM339	HFM339 Y222F F	CGGTACTTCGTCCAATGCCC TTTGACC	61
	HFM339 Y222F R	TTGGACGAAGTACCGTGAAC GGTG CAT	62
pECsf_gapS_pabABC_tu_HFM388	HFM388 Y201F F	CTCGTGTTTCGCTAATCACCC ACGCGGG	63
	HFM388 Y201F R	ATTAGCGAACACGAGTTCAT GATCGAC	64
pECsf_gapS_pabABC_tu_HFM388	HFM388 Y222F F	CGCTACTTCATCCAGTGCCC TTTGGAG	65
	HFM388 Y222F R	CTGGATGAAGTAGCGAGAA CGGGTATG	66

(3) Đưa plasmit vào tế bào chủ

Chủng *Corynebacterium glutamicum* DRHG145 (xem Đơn sáng chế Nhật Bản số 2014-523757) được biến nạp với mỗi plasmit thu được ở trên bằng công nghệ tạo lỗ xung điện (Bio-Rad Laboratories, Inc.). Dung dịch tế bào được biến nạp thu được này được cấy trên môi trường thạch LBKm và sau đó để yên ở 30 °C trong 2 ngày. Các khuẩn lạc thu được được sử dụng làm thể biến nạp.

(4) Nuôi cấy thể biến nạp

Mỗi thể biến nạp thu được ở trên được cấy vào 1 mL môi trường CGYE (chứa 50 µg/mL kanamycin sulfat) được thể hiện trong Bảng B3, và được nuôi cấy qua đêm ở 30 °C. 100 µL dung dịch nuôi cấy thu được này được cấy vào 10 mL môi trường CGXII (chứa 50 µg/mL kanamycin sulfat) được thể hiện trong Bảng B4, và được nuôi cấy ở 30 °C trong khoảng 48 giờ. Sau đó, các tế bào vi khuẩn được loại bỏ bằng cách ly tâm để thu được dịch nổi nuôi cấy. Nồng độ của axit 4-amino-3-hydroxybenzoic trong dịch nổi nuôi cấy thu được được định lượng theo phương pháp của Ví dụ tham chiếu 1. Tỷ lệ cải thiện khả năng sản xuất axit 4-amino-3-hydroxybenzoic được tính theo phương trình đưa ra dưới đây. Trong phương trình, "WT" là "thể biến nạp chứa plasmit chứa gen mã hóa enzym kiểu dại", và "MT" là "thể biến nạp chứa plasmit chứa gen mã hóa enzym đột biến và được chuẩn bị từ plasmit chứa gen mã hóa enzym kiểu dại".

(Phương trình 1)

Tỷ lệ cải thiện khả năng sản xuất = Khả năng của MT sản xuất axit 4-amino-3-hydroxybenzoic / Khả năng của WT sản xuất axit 4-amino-3-hydroxybenzoic

[Bảng B3]

Thành phần môi trường CGYE (mỗi L)

Glucoza	50 g
$(\text{NH}_4)_2\text{SO}_4$	20 g
<u>Ure</u>	5 g
KH_2PO_4	1 g
K_2HPO_4	1 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0,25 g
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	10 mg
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	10 mg
$\text{MnSO}_4 \cdot 5\text{H}_2\text{O}$	10 mg
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	1 mg
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	0,2 mg
$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$	0,02 mg
Biotin (pH7)	0,2 mg
Yeast extract	1 g

[Bảng B4]

Thành phần môi trường CGXII (mỗi L)

Glucoza	50 g
$(\text{NH}_4)_2\text{SO}_4$	20 g
<u>Ure</u>	5 g
KH_2PO_4	1 g
K_2HPO_4	1 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0,25 g

$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	10 mg
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	10 mg
$\text{MnSO}_4 \cdot 5\text{H}_2\text{O}$	10 mg
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	1 mg
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	0,2 mg
$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$	0,02 mg
Biotin (pH7)	0,2 mg
<u>Trypton</u>	10 g

(5) Kết quả

Như được thể hiện trong Bảng B5, chủng vi khuẩn chứa mỗi enzym đột biến có khả năng sản xuất axit 4-amino-3-hydroxybenzoic được cải thiện nhiều hơn so với chủng vi khuẩn chứa enzym kiểu dại.

[Bảng B5]

Hydroxylaza	Khả năng sản xuất axit 4-amino-3-hydroxybenzoic (g/L)	Tỷ lệ cải thiện khả năng sản xuất
HFM77 wt	0,073	1,00
HFM77 Y201F	0,114	1,56
HFM77 Y222F	0,100	1,37
HFM122 wt	0,134	1,00
HFM122 Y201F	0,224	1,67
HFM122 Y222F	0,256	1,90
HFM339 wt	0,016	1,00
HFM339 Y201F	0,061	3,80
HFM339 Y222F	0,139	8,74
HFM388 wt	0,033	1,00
HFM388 Y201F	0,079	2,38
HFM388 Y222F	0,230	6,95

Ví dụ C1: Sản xuất axit 4-amino-3-hydroxybenzoic

Trong ví dụ sau, PCR được thực hiện sử dụng PrimeSTAR Max Premix (Takara Bio Inc.).

(1) Chuẩn bị plasmit chứa gen mã hóa enzym kiểu dại

(a) Chuẩn bị plasmit pECsf_gapS_pabABC

Mảnh ADN chứa các gen mã hóa 4-amino-4-deoxychorismat synthaza và 4-amino-4-deoxychorismat lyaza được khuếch đại bằng PCR sử dụng bộ gen chiết xuất từ chủng *Corynebacterium glutamicum* ATCC13032 bằng phương pháp thông thường làm khuôn và sử dụng các đoạn mồi GN14_127 (SEQ ID NO: 13, TATTAATTAAATGCGCGTTTTAATTATTGATAATTATGATTC) và GN14_133 (SEQ ID NO: 14,

TTGCGGCCGCTTGTTTAAACCTCCTTACAGAAAAATGGTTGGGCG). Đoạn này được chèn vào giữa vị trí PacI và vị trí NotI của plasmit pECsf_gapS (xem Đơn sáng chế Nhật Bản số 2015-25491) để thu được plasmit pECsf_gapS_pabABC.

(b) Chuẩn bị plasmit pECsf_gapS_pabABC_HFM122

Mảnh ADN cho vectơ được tổng hợp bằng PCR sử dụng plasmit pECsf_gapS_pabABC thu được ở trên làm khuôn và sử dụng các đoạn môi pabABCcory vec R (SEQ ID NO: 15, AAATTTAAACCTCCTTTACAGAAAAATGGTTGG) và pabABCcory vec F (SEQ ID NO: 16, GGAGGTTTAAACAAGCGGCCGCGATATC). Sau đó, plasmit chứa gen (SEQ ID NO: 1) mã hóa polypeptit HFM122 có hoạt tính hydroxyl hóa axit 4-aminobenzoic được chuẩn bị bằng cách tổng hợp gen nhân tạo, và mảnh ADN cho đoạn chèn được tổng hợp bằng PCR sử dụng plasmit này làm khuôn và sử dụng các đoạn môi pECsfD HFM122 F (SEQ ID NO: 17, AGGAGGTTTAAATTTATGCGCACTCAGGTGGCTAT) và pECsfD HFM122 R (SEQ ID NO: 18, CTTGTTTAAACCTCCTTATACGAGTGGCAGTCCTA). Các sản phẩm PCR này được xử lý bằng DpnI (Takara Bio Inc.). Sau đó, các mảnh ADN tương ứng được tinh chế bằng cách sử dụng NucleoSpin Gel và PCR Clean-up (Takara Bio Inc.) và được gắn bằng cách sử dụng Kit tách dòng In-Fusion HD (Takara Bio Inc.) để xây dựng plasmit pECsf_gapS_pabABC_HFM122. Chủng *E. coli* DH5 α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch plasmit thu được. Dung dịch tế bào được cấy trên môi trường thạch LBKm (1 % Bacto Tryptone, 0,5 % chiết xuất nấm men, 1 % NaCl, 50 μ g/mL kanamycin sulfat, 1,5 % aga) và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được phải chịu phản ứng PCR sử dụng Sapphire Amp (Takara Bio Inc.) và các đoạn môi pabABC+pobA cho CPCR F (SEQ ID NO: 19, GCTATCAAAACATTCGGCACATTGGTTTTCC) và pabABC+pobA cho CPCR R (SEQ ID NO: 20, GGAAGATGCGTGATCTGATCCTTCAACTC) để chọn thể biến nạp được xác nhận là chứa mảnh ADN quan tâm. Thể biến nạp thu được được cấy vào 2 mL môi trường lỏng LBKm (1 % Bacto Tryptone, 0,5 % chiết xuất nấm men, 1 % NaCl, 50 μ g/mL kanamycin sulfat) và được nuôi cấy qua đêm ở 37 °C.

Plasmit được tinh chế từ dung dịch nuôi cấy này bằng cách sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

(c) Chuẩn bị plasmit pECsf_gapS_pabABC_tuD_HFM122

Mảnh ADN cho vector được tổng hợp bằng PCR sử dụng plasmit pECsf_gapS_pabABC_HFM122 thu được ở trên làm khuôn và sử dụng các đoạn môi pabC last R (SEQ ID NO: 21, TTACAGAAAAATGGTTGGGCGCAA) và HFM122 F (SEQ ID NO: 22, ATGCGCACTCAGGTGGCTATCG). Sau đó, một mảnh ADN (SEQ ID NO: 23, TACGTACCTGCAGGTAGCGTGTCTAGTAGGCGCGTAGGGTAAGTGGGGTAG CGGCTTGTTAGATATCTTGAAATCGGCTTTCAACAGCATTGATTTTCGATGT ATTTAGCTGGCCGTTACCCTGCGAATGTCCACAGGGTAGCTGGTAGTTTGA AAATCAACGCCGTTGCCCTTAGGATTCAGTAACTGGCACATTTTGTAATGC GCTAGATCTGTGTGCTCAGTCTTCCAGGCTGCTTATCACAGTGAAAGCAAA ACCAATTCGTGGCTGCGAAAGTCGTAGCCACCACGAAGTCCAAAGGAGGA TCTAAATTATGAATAATATAAAAGGAGGAATTAATTAA) chứa trình tự khởi động gen tuf (cg0587) (sau đây, được gọi là trình tự khởi động tu) được mang bởi chủng *Corynebacterium glutamicum* ATCC13032 được chuẩn bị bằng cách tổng hợp gen nhân tạo, và một mảnh ADN cho đoạn chèn được tổng hợp bằng PCR sử dụng đoạn này làm khuôn và sử dụng các đoạn môi pabC-Ptu F (SEQ ID NO: 24, ACCATTTTTCTGTAATACGTACCTGCAGGTAGCGTG) và PtU-HFM122 R (SEQ ID NO: 25, CACCTGAGTGCGCATTTAATTAATTCCTCCTTTTA). Các sản phẩm PCR này được xử lý bằng DpnI (Takara Bio Inc.). Sau đó, các mảnh ADN tương ứng được tinh chế bằng cách sử dụng NucleoSpin Gel và PCR Clean-up (Takara Bio Inc.) và được gắn bằng cách sử dụng Kit tách dòng In-Fusion HD (Takara Bio Inc.) để xây dựng plasmit pECsf_gapS_pabABC_tuD_HFM122. Chủng *E. coli* DH5α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch plasmit thu được. Dung dịch tế bào được cấy trên môi trường thạch LBKm và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được đã trải qua phản ứng PCR sử dụng Sapphire Amp (Takara Bio Inc.) và các đoạn môi PtU seq 1 (SEQ ID NO: 26,

GCTTGTTAGATATCTTGAAATCGGCTTTC) và pabABC+pobA cho CPCR R (SEQ ID NO: 20, GGAAGATGCGTGATCTGATCCTTCAACTC) để chọn thể biến nạp được xác nhận chứa mảnh ADN quan tâm. Thể biến nạp thu được được cấy vào 2 mL môi trường lỏng LBKm và được nuôi cấy qua đêm ở 37 °C. Plasmid được tinh chế từ dung dịch nuôi cấy này bằng cách sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

Trong plasmid được xây dựng, các gen mã hóa 4-amino-4-deoxychorismat synthaza và 4-amino-4-deoxychorismat lyaza được liên kết dưới sự kiểm soát của trình tự khởi động gap, và gen mã hóa HFM122 kiểu dại được liên kết thêm dưới sự kiểm soát của trình tự khởi động tu.

(2) Chuẩn bị plasmid chứa gen mã hóa enzym đột biến

Việc chuẩn bị plasmid chứa gen mã hóa enzym đột biến sẽ được đưa ra dưới đây bằng cách lấy ví dụ, chuẩn bị plasmid chứa gen mã hóa enzym đột biến trong đó valin ở vị trí 47 của HFM122 được thay thế bằng isoleuxin.

Plasmid pECsf_gapS_pabABC_tuD_HFM122_V47I được xây dựng bằng PCR sử dụng plasmid pECsf_gapS_pabABC_tuD_HFM122 làm khuôn và sử dụng các đoạn môi bổ sung HFM122 V47I F (SEQ ID NO: 67, GCTGGTATTCTGGAACGTATCACGGTG) và HFM122 V47I R (SEQ ID NO: 68, TTCCAGAATACCAGCCCGAACTCGGCC). Sản phẩm PCR được xử lý bằng DpnI (Takara Bio Inc.). Chúng E. coli DH5α của ECOS Competent (Nippon Gene Co., Ltd.) được biến nạp bằng dung dịch do đó đã được xử lý. Dung dịch tế bào được cấy trên môi trường thạch LBKm và sau đó để yên qua đêm ở 37 °C. Các khuẩn lạc thu được được chọn làm thể biến nạp. Thể biến nạp được cấy vào 2 mL môi trường lỏng LBKm và được nuôi cấy qua đêm ở 37 °C. Plasmid được tinh chế từ dung dịch nuôi cấy này bằng cách sử dụng NucleoSpin Plasmid EasyPure (Takara Bio Inc.).

Theo cách tương tự, plasmid chứa gen mã hóa mỗi đột biến enzym thu được bằng PCR sử dụng các đoạn môi được thể hiện trong cột "Đoạn môi" của Bảng C1 thay cho các đoạn môi HFM122 V47I F và HFM122 V47I R.

[Bảng C1]

Plasmid quan tâm	Đoạn mồi	Trình tự (5'→3')	SEQ ID NO
pECsf_gapS_pabABC_tu D_HFM122_V47S	HFM122 V47S F	GCTGGTTCTCTGGAACGTATC ACGGTG	69
	HFM122 V47S R	TTCCAGAGAACCAGCCCGAA CTCGGCC	70
pECsf_gapS_pabABC_tu D_HFM122_V47T	HFM122 V47T F	GCTGGTACACTGGAACGTAT CACGGTG	71
	HFM122 V47T R	TTCCAGTGTACCAGCCCGAA CTCGGCC	72
pECsf_gapS_pabABC_tu D_HFM122_V47C	HFM122 V47C F	GCTGGTTGTCTGGAACGTATC ACGGTG	73
	HFM122 V47C R	TTCCAGACAACCAGCCCGAA CTCGGCC	74
pECsf_gapS_pabABC_tu D_HFM122_V47M	HFM122 V47M F	GCTGGTATGCTGGAACGTAT CACGGTG	75
	HFM122 V47M R	TTCCAGCATACCAGCCCGAA CTCGGCC	76
pECsf_gapS_pabABC_tu D_HFM122_V47Q	HFM122 V47Q F	GCTGGTCAACTGGAACGTAT CACGGTG	77
	HFM122 V47Q R	TTCCAGTTGACCAGCCCGAA CTCGGCC	78
pECsf_gapS_pabABC_tu D_HFM122_H72A	HFM122 H72A F	TTGGTGGCAGCTGGCGCTAA TCTTGCG	79
	HAM122 H72A R	GCCAGCTGCCACCAAGCCCT CTCGGCG	80
pECsf_gapS_pabABC_tu D_HFM122_H72M	HFM122 H72M F	TTGGTGATGGCTGGCGCTAAT CTTGCG	81
	HFM122 H72M R	GCCAGCCATCACCAAGCCCT CTCGGCG	82
pECsf_gapS_pabABC_tu	HFM122	TTTGCCATGGCGTCGATGCGC	83

D_HFM122_L210M	L210M F	TCACCG	
	HFM122 L210M R	CGACGCCATGGCAAAACCGC GATCATG	84
pECsf_gapS_pabABC_tu D_HFM122_T294A	HFM122 T294A F	CCACCCGCAGGAGCGAAAGG GATGAAC	85
	HFM122 T294A R	CGCTCCTGCGGGTGGAAACGA TATGAGC	86
pECsf_gapS_pabABC_tu D_HFM122_T294G	HFM122 T294G F	CCACCCGGTGGAGCGAAAGG GATGAAC	87
	HFM122 T294G R	CGCTCCACCGGGTGGAAACGA TATGAGC	88
pECsf_gapS_pabABC_tu D_HFM122_T294C	HFM122 T294C F	CCACCCTGTGGAGCGAAAGG GATGAAC	89
	HFM122 T294C R	CGCTCCACAGGGTGGAAACGA TATGAGC	90
pECsf_gapS_pabABC_tu D_HFM122_T294S	HFM122 T294S F	CCACCCTCTGGAGCGAAAGG GATGAAC	91
	HFM122 T294S R	CGCTCCAGAGGGTGGAAACGA TATGAGC	92
pECsf_gapS_pabABC_tu D_HFM122_Y385V	HFM122 Y385V F	GAGAACGTTGTAGGACTGCC ACTCGTA	93
	HFM122 Y385V R	TCCTACAACGTTCTCCGCCAG GGTGAC	94
pECsf_gapS_pabABC_tu D_HFM122_Y385L	HFM122 Y385L F	GAGAACCTCGTAGGACTGCC ACTCGTA	95
	HFM122 Y385L R	TCCTACGAGGTTCTCCGCCAG GGTGAC	96
pECsf_gapS_pabABC_tu D_HFM122_Y385I	HFM122 Y385I F	GAGAACATTGTAGGACTGCC ACTCGTA	97
	HFM122 Y385I R	TCCTACAATGTTCTCCGCCAG GGTGAC	98
pECsf_gapS_pabABC_tu	HFM122	GAGAACATGGTAGGACTGCC	99

D_HFM122_Y385M	Y385M F	ACTCGTA	
	HFM122 Y385M R	TCCTACCATGTTCTCCGCCAG GGTGAC	100

(3) Đưa plasmit vào tế bào chủ

Chủng *Corynebacterium glutamicum* DRHG145 (xem Đơn sáng chế Nhật Bản số 2014-523757) được biến nạp bằng mỗi plasmit thu được ở trên bằng công nghệ tạo lỗ xung điện (Bio-Rad Laboratories, Inc.). Dung dịch tế bào được biến nạp thu được này được cấy trên môi trường thạch LBKm và sau đó để yên ở 30 °C trong 2 ngày. Các khuẩn lạc thu được được sử dụng làm thể biến nạp.

(4) Nuôi cấy thể biến nạp

Mỗi thể biến nạp thu được ở trên được cấy vào 1 mL môi trường CGYE (chứa 50 µg/mL kanamycin sulfat) được thể hiện trong Bảng C2, và được nuôi cấy qua đêm ở 30 °C. 100 µL dung dịch nuôi cấy thu được được cấy vào 10 mL môi trường CGXII (chứa 50 µg/mL kanamycin sulfat) được thể hiện trong Bảng C3, và được nuôi cấy ở 30 °C trong khoảng 48 giờ. Sau đó, các tế bào vi khuẩn được loại bỏ bằng cách ly tâm để thu được dịch nổi nuôi cấy. Nồng độ của axit 4-amino-3-hydroxybenzoic trong dịch nổi nuôi cấy thu được được định lượng theo phương pháp của Ví dụ tham chiếu 1. Tỷ lệ cải thiện khả năng sản xuất axit 4-amino-3-hydroxybenzoic được tính theo phương trình đưa ra dưới đây. Trong bối cảnh này, "WT" là "thể biến nạp chứa plasmit chứa gen mã hóa enzym kiểu dại", và "MT" là "thể biến nạp chứa plasmit chứa gen mã hóa enzym đột biến và được chuẩn bị từ plasmit chứa gen mã hóa enzym kiểu dại".

(Phương trình 1)

Tỷ lệ cải thiện khả năng sản xuất = Khả năng của MT sản xuất axit 4-amino-3-hydroxybenzoic / Khả năng của WT sản xuất axit 4-amino-3-hydroxybenzoic

[Bảng C2]

Thành phần môi trường CGYE (mỗi L)

Glucoza	50 g
$(\text{NH}_4)_2\text{SO}_4$	20 g
<u>Ure</u>	5 g
KH_2PO_4	1 g
K_2HPO_4	1 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0,25 g
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	10 mg
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	10 mg
$\text{MnSO}_4 \cdot 5\text{H}_2\text{O}$	10 mg
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	1 mg
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	0,2 mg
$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$	0,02 mg
Biotin (pH7)	0,2 mg
Chiết xuất nấm men	1 g

[Bảng C3]

Thành phần môi trường CGXII (mỗi L)

Glucoza	50 g
$(\text{NH}_4)_2\text{SO}_4$	20 g
<u>Ure</u>	5 g
KH_2PO_4	1 g
K_2HPO_4	1 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	0,25 g

$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	10 mg
$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	10 mg
$\text{MnSO}_4 \cdot 5\text{H}_2\text{O}$	10 mg
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	1 mg
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	0,2 mg
$\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$	0,02 mg
Biotin (pH7)	0,2 mg
<u>Trypton</u>	10 g

(5) Kết quả

Như được thể hiện trong Bảng C4, chủng vi khuẩn chứa mỗi enzym đột biến có khả năng sản xuất axit 4-amino-3-hydroxybenzoic được cải thiện nhiều hơn so với chủng vi khuẩn chứa enzym kiểu dại.

[Bảng C4]

Hydroxylaza	Khả năng sản xuất axit 4-amino-3-hydroxybenzoic (g/L)	Tỷ lệ cải thiện khả năng sản xuất
HFM122 wt	0,134	1,00
HFM122 V47I	0,306	2,28
HFM122 V47S	0,163	1,21
HFM122 V47T	0,202	1,50
HFM122 V47C	0,174	1,30
HFM122 V47M	0,290	2,16
HFM122 V47Q	0,254	1,89
HFM122 H72A	0,164	1,22
HFM122 H72M	0,216	1,61
HFM122 L210M	0,173	1,29
HFM122 T294A	0,160	1,19
HFM122 T294G	0,187	1,39
HFM122 T294C	0,190	1,41
HFM122 T294S	0,378	2,81
HFM122 Y385V	0,157	1,17
HFM122 Y385L	0,168	1,25
HFM122 Y385I	0,186	1,38
HFM122 Y385M	0,167	1,24

Ví dụ tham chiếu 1: Định lượng axit 4-amino-3-hydroxybenzoic

Axit 4-Amino-3-hydroxybenzoic được định lượng bằng HPLC. Dung dịch phản ứng đưa ra để phân tích HPLC được pha loãng thích hợp với 0,1 % axit phosphoric.

Sau đó, chất không hòa tan được loại bỏ bằng cách sử dụng các tấm lọc AcroPrep 96 giếng (màng GHP 0,2 μm , Nihon Pall Ltd.).

Thiết bị HPLC được sử dụng là Chromaster (Hitachi High-Tech Science Corp.). Cột phân tích được sử dụng là cột L- ODS (4,6 mm I.D. $\times$ 150 mm, Viện Nghiên cứu và Đánh giá Hóa chất, Nhật Bản). Dung môi rửa giải A là dung dịch axit phosphoric 0,1 % của kali dihydro photphat 0,1 M, và dung môi rửa giải B là 70 % metanol. Rửa giải gradient được thực hiện trong các điều kiện liên quan đến lưu lượng dòng chảy 1,0 mL/phút và nhiệt độ cột là 40 °C. Một máy dò UV (bước sóng dò: 280 nm) được sử dụng để dò axit 4-amino-3-hydroxybenzoic. Đường chuẩn nồng độ được chuẩn bị bằng khuôn chuẩn [axit 4-amino-3-hydroxybenzoic (mã nhà phân phối A1194, Tokyo Chemical Industry Co., Ltd.)]. Axit 4-Amino-3-hydroxybenzoic được định lượng trên cơ sở đường chuẩn nồng độ.

YÊU CẦU BẢO HỘ

1. Polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, polypeptit được chọn từ nhóm bao gồm A) đến C) sau đây:

A) polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2, 4, 6, 8, 10 hoặc 12, hoặc trình tự axit amin có độ tương đồng ít nhất 90% với chúng, và có gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó là leuxin;

B) polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2, 4, 6 hoặc 8, hoặc trình tự axit amin có độ tương đồng ít nhất 90% với chúng, và có gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 là phenylalanin; và

C) polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic, gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 là axit amin sau:

(a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,

(b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,

(c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,

(d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và

(e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

2. Phương pháp sản xuất polypeptit đột biến có hoạt tính hydroxyl hóa axit 4-aminobenzoic, bao gồm thay thế gốc axit amin, thay thế được chọn từ nhóm bao gồm A') đến C') sau đây:

A') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2, 4, 6, 8, 10 hoặc 12, hoặc trình tự axit amin có độ tương đồng ít nhất 90% với chúng, và

có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó bằng leuxin;

B') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2, 4, 6 hoặc 8, hoặc trình tự axit amin có độ tương đồng ít nhất 90% với chúng, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 bằng phenylalanin; và

C') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 bằng axit amin sau đây:

(a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,

(b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,

(c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,

(d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và

(e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

3. Phương pháp cải thiện hoạt tính hydroxyl hóa axit 4-aminobenzoic, bao gồm thay thế gốc axit amin, thay thế được chọn từ nhóm bao gồm A') đến C') sau đây:

A') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2, 4, 6, 8, 10 hoặc 12, hoặc trình tự axit amin có độ tương đồng ít nhất 90% với chúng, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 47 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc vị trí tương ứng với nó bằng leuxin;

B') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2, 4, 6, hoặc 8, hoặc trình tự axit amin có độ tương đồng ít nhất 90% với chúng, và có hoạt

tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 201 hoặc 222 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 201 hoặc 222 bằng phenylalanin; và

C') trong polypeptit gồm trình tự axit amin được thể hiện bởi SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng ít nhất 90% với nó, và có hoạt tính hydroxyl hóa axit 4-aminobenzoic, thay thế gốc axit amin ở vị trí 47, 72, 210, 294 hoặc 385 của trình tự axit amin được thể hiện bởi SEQ ID NO: 2, hoặc vị trí tương ứng với vị trí 47, 72, 210, 294 hoặc 385 bằng axit amin sau đây:

(a) vị trí 47 hoặc vị trí tương ứng với nó: isoleuxin, serin, threonin, xystein, methionin, hoặc glutamin,

(b) vị trí 72 hoặc vị trí tương ứng với nó: alanin hoặc methionin,

(c) vị trí 210 hoặc vị trí tương ứng với nó: methionin,

(d) vị trí 294 hoặc vị trí tương ứng với nó: alanin, glyxin, xystein, hoặc serin, và

(e) vị trí 385 hoặc vị trí tương ứng với nó: valin, leuxin, isoleuxin, hoặc methionin.

4. Phương pháp theo điểm 2 hoặc 3, trong đó sự thay thế gốc axit amin trong A') là sự thay thế valin bằng leuxin, và sự thay thế gốc axit amin trong B') là sự thay thế tyrosin bằng phenylalanin.

5. Polynucleotit mã hóa polypeptit theo điểm 1.

6. Vectơ hoặc mảnh ADN bao gồm polynucleotit theo điểm 5.

7. Tế bào được biến nạp bao gồm vectơ hoặc mảnh ADN theo điểm 6.

8. Tế bào được biến nạp theo điểm 7, trong đó tế bào được biến nạp này thu được từ *E. coli* hoặc vi khuẩn thuộc chi *Corynebacterium*.

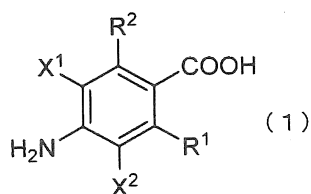
9. Tế bào được biến nạp theo điểm 7 hoặc 8, trong đó tế bào được biến nạp này là vi sinh vật có khả năng cung cấp axit 4-aminobenzoic.

10. Phương pháp sản xuất axit 4-amino-3-hydroxybenzoic, trong đó phương pháp này bao gồm bước nuôi cấy tế bào được biến nạp theo điểm bất kỳ trong số các điểm từ 7 đến 9 trong sự có mặt của axit 4-aminobenzoic.

11. Phương pháp theo điểm 10, trong đó phương pháp này còn bao gồm bước thu gom axit 4-amino-3-hydroxybenzoic từ môi trường nuôi cấy.

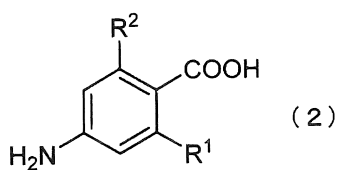
12. Phương pháp theo điểm 10 hoặc 11, trong đó:

axit 4-amino-3-hydroxybenzoic là các dẫn xuất của axit 4-amino-3-hydroxybenzoic có công thức (1) sau:



trong đó R^1 là nguyên tử hydro, nhóm hydroxy, nhóm metoxy, nhóm amin, nguyên tử flo, nguyên tử clo, nguyên tử brom, nguyên tử iốt, nhóm cacboxy, nhóm metyl, hoặc nhóm etyl, R^2 là nguyên tử hydro, nhóm hydroxy, nhóm metoxy, nhóm amin, nguyên tử flo, nguyên tử clo, nguyên tử brom, nguyên tử iốt, nhóm cacboxy, nhóm metyl, hoặc nhóm etyl, và X^1 và X^2 mỗi loại là nguyên tử hydro hoặc nhóm hydroxy, ít nhất một trong số đó là nhóm hydroxy, và

axit 4-aminobenzoic là các dẫn xuất của axit 4-aminobenzoic có công thức (2) sau:



trong đó R^1 là nguyên tử hydro, nhóm hydroxy, nhóm metoxy, nhóm amin, nguyên tử flo, nguyên tử clo, nguyên tử brom, nguyên tử iốt, nhóm cacboxy, nhóm metyl, hoặc nhóm etyl, và R^2 là nguyên tử hydro, nhóm hydroxy, nhóm metoxy, nhóm amin, nguyên tử flo, nguyên tử clo, nguyên tử brom, nguyên tử iốt, nhóm cacboxy, nhóm metyl, hoặc nhóm etyl.

Danh mục trình tự

<110> KAO CORPORATION

<120> Polypeptit có hoạt tính hydroxyl hóa axit 4-aminobenzoic và các phương pháp sử dụng chúng

<130> KS1681

<150> JP 2019-203523

<151> 2019-11-08

<150> JP 2019-233484

<151> 2019-12-24

<150> JP 2019-233485

<151> 2019-12-24

<160> 100

<170> PatentIn phiên bản 3.5

<210> 1

<211> 1176

<212> ADN

<213> Caulobacter vibrioides

<220>

<221> CDS

<222> (1)..(1173)

<223> Oligonucleotit được tối ưu hóa cho codon

<400> 1

atg cgc act cag gtg gct atc gta gga gca ggc cca gct ggc ctg ttc 48

Met Arg Thr Gln Val Ala Ile Val Gly Ala Gly Pro Ala Gly Leu Phe

1 5 10 15

ttg ggc cat ctc ctc cgt caa gct ggt gtg gac gtc gtg att ctg gaa 96

Leu Gly His Leu Leu Arg Gln Ala Gly Val Asp Val Val Ile Leu Glu

20 25 30

cgc aaa gac cgt gct tat gtc gaa ggc cga gtt cgg gct ggt gtc ctg 144
 Arg Lys Asp Arg Ala Tyr Val Glu Gly Arg Val Arg Ala Gly Val Leu
 35 40 45
 gaa cgt atc acg gtg gag ctg atg gag cgt ctt ggt gtg gat gag cga 192
 Glu Arg Ile Thr Val Glu Leu Met Glu Arg Leu Gly Val Asp Glu Arg
 50 55 60
 atg cgc cga gag ggc ttg gtg cat gct ggc gct aat ctt gcg tct gat 240
 Met Arg Arg Glu Gly Leu Val His Ala Gly Ala Asn Leu Ala Ser Asp
 65 70 75 80
 ggc gag atg ttc cgt atc gac atg gca gag ctc acg ggt ggt tcc acc 288
 Gly Glu Met Phe Arg Ile Asp Met Ala Glu Leu Thr Gly Gly Ser Thr
 85 90 95
 gtc atg gtt tac ggc caa cag gag gtg atg aag gac ctg ttt gat gca 336
 Val Met Val Tyr Gly Gln Gln Glu Val Met Lys Asp Leu Phe Asp Ala
 100 105 110
 gca gag cag cgc gat ctg cga att gtc ttt gac gcc gat gca gtg cgt 384
 Ala Glu Gln Arg Asp Leu Arg Ile Val Phe Asp Ala Asp Ala Val Arg
 115 120 125
 ctg cac gat gtg gaa ggc gaa cgt cct cac atc acc tgg cgc aaa gac 432
 Leu His Asp Val Glu Gly Glu Arg Pro His Ile Thr Trp Arg Lys Asp
 130 135 140
 ggg gca gaa cac cgc ctg gac tgc gat ttc att gcc ggc tgc gac ggc 480
 Gly Ala Glu His Arg Leu Asp Cys Asp Phe Ile Ala Gly Cys Asp Gly
 145 150 155 160
 tac cac gga gtt tct cgt gcg acc att ccc gat aag gtt ctc aag acc 528
 Tyr His Gly Val Ser Arg Ala Thr Ile Pro Asp Lys Val Leu Lys Thr
 165 170 175
 ttc gaa cgg gtg tat ccc ttt ggg tgg ttg gga atc ctg gct gaa gca 576
 Phe Glu Arg Val Tyr Pro Phe Gly Trp Leu Gly Ile Leu Ala Glu Ala

180	185	190	
cct ccg tgt gac cac gag ttg atc tac tcg aac cat gat cgc ggt ttt			624
Pro Pro Cys Asp His Glu Leu Ile Tyr Ser Asn His Asp Arg Gly Phe			
195	200	205	
gcc ctg gcg tcg atg cgc tca ccg aca cgc tcc cgc tat tac gtg cag			672
Ala Leu Ala Ser Met Arg Ser Pro Thr Arg Ser Arg Tyr Tyr Val Gln			
210	215	220	
tgc tca ctc gac gat cgc ctc gag gat tgg tcc gat gaa cgg ttc tgg			720
Cys Ser Leu Asp Asp Arg Leu Glu Asp Trp Ser Asp Glu Arg Phe Trp			
225	230	235	240
gat gaa gtt tcg gtt cgc ctg gga ccg gaa gca gcc gct cgg atc gtt			768
Asp Glu Val Ser Val Arg Leu Gly Pro Glu Ala Ala Ala Arg Ile Val			
245	250	255	
cgc gca cct tcc ttc gag aag agc att gcc cca ctt cgc tcc ttc gtt			816
Arg Ala Pro Ser Phe Glu Lys Ser Ile Ala Pro Leu Arg Ser Phe Val			
260	265	270	
tcc gag cct atg cgg tat ggc cgc ctt ttc ctc gcg ggt gat gcg gct			864
Ser Glu Pro Met Arg Tyr Gly Arg Leu Phe Leu Ala Gly Asp Ala Ala			
275	280	285	
cat atc gtt cca ccc act gga gcg aaa ggg atg aac ttg gcc gta tca			912
His Ile Val Pro Pro Thr Gly Ala Lys Gly Met Asn Leu Ala Val Ser			
290	295	300	
gac gtc atc atg ctg tcc gaa gcc ctg gtc gaa cac tac cac gaa cgc			960
Asp Val Ile Met Leu Ser Glu Ala Leu Val Glu His Tyr His Glu Arg			
305	310	315	320
tct tcc gct ggt atc gat ggt tac agc gca cgt gca ctt gcc cgc gtc			1008
Ser Ser Ala Gly Ile Asp Gly Tyr Ser Ala Arg Ala Leu Ala Arg Val			
325	330	335	
tgg aag gcg gag cgt ttc agc tgg tgg ttt acc tcc ctt act cac cgc			1056

Trp Lys Ala Glu Arg Phe Ser Trp Trp Phe Thr Ser Leu Thr His Arg
 340 345 350
 ttc cca gac cag gac ggc ttc gac cgc aag atg caa gtc gcc gaa ttg 1104
 Phe Pro Asp Gln Asp Gly Phe Asp Arg Lys Met Gln Val Ala Glu Leu
 355 360 365
 gca tac atc aag ggt tct cgc gct gcc cag gtc acc ctg gcg gag aac 1152
 Ala Tyr Ile Lys Gly Ser Arg Ala Ala Gln Val Thr Leu Ala Glu Asn
 370 375 380
 tac gta gga ctg cca ctc gta taa 1176
 Tyr Val Gly Leu Pro Leu Val
 385 390
 <210> 2
 <211> 391
 <212> PRT
 <213> *Caulobacter vibrioides*
 <400> 2
 Met Arg Thr Gln Val Ala Ile Val Gly Ala Gly Pro Ala Gly Leu Phe
 1 5 10 15
 Leu Gly His Leu Leu Arg Gln Ala Gly Val Asp Val Val Ile Leu Glu
 20 25 30
 Arg Lys Asp Arg Ala Tyr Val Glu Gly Arg Val Arg Ala Gly Val Leu
 35 40 45
 Glu Arg Ile Thr Val Glu Leu Met Glu Arg Leu Gly Val Asp Glu Arg
 50 55 60
 Met Arg Arg Glu Gly Leu Val His Ala Gly Ala Asn Leu Ala Ser Asp
 65 70 75 80
 Gly Glu Met Phe Arg Ile Asp Met Ala Glu Leu Thr Gly Gly Ser Thr
 85 90 95
 Val Met Val Tyr Gly Gln Gln Glu Val Met Lys Asp Leu Phe Asp Ala

```

      100              105              110
Ala Glu Gln Arg Asp Leu Arg Ile Val Phe Asp Ala Asp Ala Val Arg

      115              120              125
Leu His Asp Val Glu Gly Glu Arg Pro His Ile Thr Trp Arg Lys Asp

      130              135              140
Gly Ala Glu His Arg Leu Asp Cys Asp Phe Ile Ala Gly Cys Asp Gly

      145              150              155              160
Tyr His Gly Val Ser Arg Ala Thr Ile Pro Asp Lys Val Leu Lys Thr

      165              170              175
Phe Glu Arg Val Tyr Pro Phe Gly Trp Leu Gly Ile Leu Ala Glu Ala

      180              185              190
Pro Pro Cys Asp His Glu Leu Ile Tyr Ser Asn His Asp Arg Gly Phe

      195              200              205
Ala Leu Ala Ser Met Arg Ser Pro Thr Arg Ser Arg Tyr Tyr Val Gln

      210              215              220
Cys Ser Leu Asp Asp Arg Leu Glu Asp Trp Ser Asp Glu Arg Phe Trp

      225              230              235              240
Asp Glu Val Ser Val Arg Leu Gly Pro Glu Ala Ala Ala Arg Ile Val

      245              250              255
Arg Ala Pro Ser Phe Glu Lys Ser Ile Ala Pro Leu Arg Ser Phe Val

      260              265              270
Ser Glu Pro Met Arg Tyr Gly Arg Leu Phe Leu Ala Gly Asp Ala Ala

      275              280              285
His Ile Val Pro Pro Thr Gly Ala Lys Gly Met Asn Leu Ala Val Ser

      290              295              300
Asp Val Ile Met Leu Ser Glu Ala Leu Val Glu His Tyr His Glu Arg

      305              310              315              320
Ser Ser Ala Gly Ile Asp Gly Tyr Ser Ala Arg Ala Leu Ala Arg Val

      325              330              335

```

Trp Lys Ala Glu Arg Phe Ser Trp Trp Phe Thr Ser Leu Thr His Arg

340

345

350

Phe Pro Asp Gln Asp Gly Phe Asp Arg Lys Met Gln Val Ala Glu Leu

355

360

365

Ala Tyr Ile Lys Gly Ser Arg Ala Ala Gln Val Thr Leu Ala Glu Asn

370

375

380

Tyr Val Gly Leu Pro Leu Val

385

390

<210> 3

<211> 1173

<212> ADN

<213> Sinorhizobium meliloti

<220>

<221> CDS

<222> (1)..(1170)

<223> Oligonucleotit được tối ưu hóa cho codon

<400> 3

atg cgc acc caa gtg gtc atc atc ggc tca gga ccg tct ggc ctt ctt 48

Met Arg Thr Gln Val Val Ile Ile Gly Ser Gly Pro Ser Gly Leu Leu

1

5

10

15

ctg gga cag ctt ctg acc gag gca ggg atc gca aac gtc atc ctc gat 96

Leu Gly Gln Leu Leu Thr Glu Ala Gly Ile Ala Asn Val Ile Leu Asp

20

25

30

cgc gct acc aag gcc cac att ctc ggg cga gtt cgc gct gga gtg ttg 144

Arg Ala Thr Lys Ala His Ile Leu Gly Arg Val Arg Ala Gly Val Leu

35

40

45

gaa cag ggc acc gtt cgc ctt atg gaa gag gct ggt tgt ggt gcg cga 192

Glu Gln Gly Thr Val Arg Leu Met Glu Glu Ala Gly Cys Gly Ala Arg

50

55

60

atg cac gca gaa gga ctg cca cac gac ggc ttt tcg ctg gca ttc gac 240
 Met His Ala Glu Gly Leu Pro His Asp Gly Phe Ser Leu Ala Phe Asp
 65 70 75 80
 ggt cgg gat cac cgc att gac ctg ttc ggc ctg act ggt ggc cgt cgc 288
 Gly Arg Asp His Arg Ile Asp Leu Phe Gly Leu Thr Gly Gly Arg Arg
 85 90 95
 gta atg atc tat ggt cag acg gag ctg aca cgg gac ctg atg gat cat 336
 Val Met Ile Tyr Gly Gln Thr Glu Leu Thr Arg Asp Leu Met Asp His
 100 105 110
 cgc gag cgg gtt ggt gca ctc tcc atc tac gaa gcg gct aac gtc atg 384
 Arg Glu Arg Val Gly Ala Leu Ser Ile Tyr Glu Ala Ala Asn Val Met
 115 120 125
 ccc cgc gat ttc gat gga cgg aca cct cac gtt gcg tat gag aag gac 432
 Pro Arg Asp Phe Asp Gly Arg Thr Pro His Val Ala Tyr Glu Lys Asp
 130 135 140
 ggt att gcg caa cgc att gac tgc gac ttc atc gcc ggc tgt gac ggc 480
 Gly Ile Ala Gln Arg Ile Asp Cys Asp Phe Ile Ala Gly Cys Asp Gly
 145 150 155 160
 ttc cat ggc gtg tcc cgt cgt tcc ctg cca gag aaa gcc atc cgg aat 528
 Phe His Gly Val Ser Arg Arg Ser Leu Pro Glu Lys Ala Ile Arg Asn
 165 170 175
 ttc gag aag atc tac cca ttt gga tgg ctc ggc atc ctg gca gat gtc 576
 Phe Glu Lys Ile Tyr Pro Phe Gly Trp Leu Gly Ile Leu Ala Asp Val
 180 185 190
 cca ccg gtc gat cat gaa ctc gtg tac gct aat cac cca cgc ggg ttt 624
 Pro Pro Val Asp His Glu Leu Val Tyr Ala Asn His Pro Arg Gly Phe
 195 200 205
 gca ttg tgc tct atg cgt tcc cat acc cgt tct cgc tac tac atc cag 672
 Ala Leu Cys Ser Met Arg Ser His Thr Arg Ser Arg Tyr Tyr Ile Gln

210	215	220	
tgc cct ttg gag gaa aag atc gaa gat tgg gat gat cag cgc ttc tgg			720
Cys Pro Leu Glu Glu Lys Ile Glu Asp Trp Asp Asp Gln Arg Phe Trp			
225	230	235	240
gat gaa ctc cgc cgt cgt ctg cct gcg cac cat gcg gag cgc gtg gta			768
Asp Glu Leu Arg Arg Arg Leu Pro Ala His His Ala Glu Arg Val Val			
245	250	255	
acc ggt ccg agc ttc gag aag tcc att gca ccc ttg cgt tcg ttc gtt			816
Thr Gly Pro Ser Phe Glu Lys Ser Ile Ala Pro Leu Arg Ser Phe Val			
260	265	270	
gcc gaa cca atg cgc ttc aac cgc ctg ttc ttg gcg ggc gat gct gcc			864
Ala Glu Pro Met Arg Phe Asn Arg Leu Phe Leu Ala Gly Asp Ala Ala			
275	280	285	
cac att gtc cca cct acc ggt gct aaa ggc ctt aac ctc gct gct tcc			912
His Ile Val Pro Pro Thr Gly Ala Lys Gly Leu Asn Leu Ala Ala Ser			
290	295	300	
gac gtc cac tac ctg ttt gag ggg ttg ctc gaa cac tac cag gat cga			960
Asp Val His Tyr Leu Phe Glu Gly Leu Leu Glu His Tyr Gln Asp Arg			
305	310	315	320
agc aac gca ggc atc gac gcc tat tcc gct cgc gca ctg gct cgt gtg			1008
Ser Asn Ala Gly Ile Asp Ala Tyr Ser Ala Arg Ala Leu Ala Arg Val			
325	330	335	
tgg aaa gcc gtt cgc ttc agc tgg tgg atg acg act atg ctt cac cgt			1056
Trp Lys Ala Val Arg Phe Ser Trp Trp Met Thr Thr Met Leu His Arg			
340	345	350	
ttt ccc gaa acc tcc gac ttt gac cag cgc att caa gag gcc gaa ctg			1104
Phe Pro Glu Thr Ser Asp Phe Asp Gln Arg Ile Gln Glu Ala Glu Leu			
355	360	365	
gac tat ctc acc cac tca cga gct gcc gca act gca ctt gcg gaa aac			1152

Asp Tyr Leu Thr His Ser Arg Ala Ala Ala Thr Ala Leu Ala Glu Asn

370 375 380

tac gtg ggt ctg ccg ttc taa 1173

Tyr Val Gly Leu Pro Phe

385 390

<210> 4

<211> 390

<212> PRT

<213> Sinorhizobium meliloti

<400> 4

Met Arg Thr Gln Val Val Ile Ile Gly Ser Gly Pro Ser Gly Leu Leu

1 5 10 15

Leu Gly Gln Leu Leu Thr Glu Ala Gly Ile Ala Asn Val Ile Leu Asp

20 25 30

Arg Ala Thr Lys Ala His Ile Leu Gly Arg Val Arg Ala Gly Val Leu

35 40 45

Glu Gln Gly Thr Val Arg Leu Met Glu Glu Ala Gly Cys Gly Ala Arg

50 55 60

Met His Ala Glu Gly Leu Pro His Asp Gly Phe Ser Leu Ala Phe Asp

65 70 75 80

Gly Arg Asp His Arg Ile Asp Leu Phe Gly Leu Thr Gly Gly Arg Arg

85 90 95

Val Met Ile Tyr Gly Gln Thr Glu Leu Thr Arg Asp Leu Met Asp His

100 105 110

Arg Glu Arg Val Gly Ala Leu Ser Ile Tyr Glu Ala Ala Asn Val Met

115 120 125

Pro Arg Asp Phe Asp Gly Arg Thr Pro His Val Ala Tyr Glu Lys Asp

130 135 140

Gly Ile Ala Gln Arg Ile Asp Cys Asp Phe Ile Ala Gly Cys Asp Gly

145 150 155 160
Phe His Gly Val Ser Arg Arg Ser Leu Pro Glu Lys Ala Ile Arg Asn
 165 170 175
Phe Glu Lys Ile Tyr Pro Phe Gly Trp Leu Gly Ile Leu Ala Asp Val
 180 185 190
Pro Pro Val Asp His Glu Leu Val Tyr Ala Asn His Pro Arg Gly Phe
 195 200 205
Ala Leu Cys Ser Met Arg Ser His Thr Arg Ser Arg Tyr Tyr Ile Gln
 210 215 220
Cys Pro Leu Glu Glu Lys Ile Glu Asp Trp Asp Asp Gln Arg Phe Trp
225 230 235 240
Asp Glu Leu Arg Arg Arg Leu Pro Ala His His Ala Glu Arg Val Val
 245 250 255
Thr Gly Pro Ser Phe Glu Lys Ser Ile Ala Pro Leu Arg Ser Phe Val
 260 265 270
Ala Glu Pro Met Arg Phe Asn Arg Leu Phe Leu Ala Gly Asp Ala Ala
 275 280 285
His Ile Val Pro Pro Thr Gly Ala Lys Gly Leu Asn Leu Ala Ala Ser
 290 295 300
Asp Val His Tyr Leu Phe Glu Gly Leu Leu Glu His Tyr Gln Asp Arg
305 310 315 320
Ser Asn Ala Gly Ile Asp Ala Tyr Ser Ala Arg Ala Leu Ala Arg Val
 325 330 335
Trp Lys Ala Val Arg Phe Ser Trp Trp Met Thr Thr Met Leu His Arg
 340 345 350
Phe Pro Glu Thr Ser Asp Phe Asp Gln Arg Ile Gln Glu Ala Glu Leu
 355 360 365
Asp Tyr Leu Thr His Ser Arg Ala Ala Ala Thr Ala Leu Ala Glu Asn
 370 375 380

Tyr Val Gly Leu Pro Phe

385 390

<210> 5

<211> 1173

<212> ADN

<213> Rhodopseudomonas palustris

<220>

<221> CDS

<222> (1)..(1170)

<223> Oligonucleotit được tối ưu hóa cho codon

<400> 5

atg cgc act cag gtg gca atc att ggt gcc ggt cca tcc ggt ctg ctt 48

Met Arg Thr Gln Val Ala Ile Ile Gly Ala Gly Pro Ser Gly Leu Leu

1 5 10 15

ctc gga cag ctc ctt cac aag tac ggg atc gac gcc gtc atc gtt gag 96

Leu Gly Gln Leu Leu His Lys Tyr Gly Ile Asp Ala Val Ile Val Glu

20 25 30

cgc aaa gat ccc gat tac gtg ctt tcc cgc atc cgt gca ggc gtt ctg 144

Arg Lys Asp Pro Asp Tyr Val Leu Ser Arg Ile Arg Ala Gly Val Leu

35 40 45

gag cag ggc atg gtt gac ctc ctg gac gaa gct ggc gta agc gct cgg 192

Glu Gln Gly Met Val Asp Leu Leu Asp Glu Ala Gly Val Ser Ala Arg

50 55 60

ttg cat cag gaa gcg ctt gtt cac ggt ggt ttc gaa atc gcg ttc gca 240

Leu His Gln Glu Ala Leu Val His Gly Gly Phe Glu Ile Ala Phe Ala

65 70 75 80

ggc caa cgg cat ccc att gat ctg cgt ggt gca acc gga ggc aag tct 288

Gly Gln Arg His Pro Ile Asp Leu Arg Gly Ala Thr Gly Gly Lys Ser

85 90 95

gtc acc gtc tat ggt cag acc gag gta acc cga gat ctc atg gag gct 336
 Val Thr Val Tyr Gly Gln Thr Glu Val Thr Arg Asp Leu Met Glu Ala
 100 105 110
 cga tct gca gct ggc ctt acg acg atc tat gac gct gcg gat gtt tcg 384
 Arg Ser Ala Ala Gly Leu Thr Thr Ile Tyr Asp Ala Ala Asp Val Ser
 115 120 125
 ctg cat gac ttt gaa ggc gca cac cct aag gtg cgt tac gtc aaa gac 432
 Leu His Asp Phe Glu Gly Ala His Pro Lys Val Arg Tyr Val Lys Asp
 130 135 140
 ggc acc act cgc gag atc gtg tgc gat ttc att gcg gga tgt gac ggg 480
 Gly Thr Thr Arg Glu Ile Val Cys Asp Phe Ile Ala Gly Cys Asp Gly
 145 150 155 160
 ttc cac gga att tcg cgc caa tct gtg cca gcg tct gcc gtt cag tcc 528
 Phe His Gly Ile Ser Arg Gln Ser Val Pro Ala Ser Ala Val Gln Ser
 165 170 175
 ttc gaa cgc gtg tat ccg ttc ggc tgg ttg gga ctt ctg tcc gat acc 576
 Phe Glu Arg Val Tyr Pro Phe Gly Trp Leu Gly Leu Leu Ser Asp Thr
 180 185 190
 cct ccc gtc agc cca gaa ctg atc tac gtc aac cac gac cga ggc ttt 624
 Pro Pro Val Ser Pro Glu Leu Ile Tyr Val Asn His Asp Arg Gly Phe
 195 200 205
 gct ctc tgc tcg atg cgc tcc atg cac cgt tca cgg tac tac gtc caa 672
 Ala Leu Cys Ser Met Arg Ser Met His Arg Ser Arg Tyr Tyr Val Gln
 210 215 220
 tgc cct ttg acc gac gat gtc gct gat tgg tct gat gat cgc ttc tgg 720
 Cys Pro Leu Thr Asp Asp Val Ala Asp Trp Ser Asp Asp Arg Phe Trp
 225 230 235 240
 gac gaa ctg aag tcc cgc ctc gac ccg gaa aca gcg ggg aaa ctc gtg 768
 Asp Glu Leu Lys Ser Arg Leu Asp Pro Glu Thr Ala Gly Lys Leu Val

245	250	255	
aca ggc ccg tca atc gag aag tca atc gca cca ctg cgt tcc ttt gtc			816
Thr Gly Pro Ser Ile Glu Lys Ser Ile Ala Pro Leu Arg Ser Phe Val			
260	265	270	
gct gaa cca atg cgc ttt ggt cgc ctt ttc ctg gcc ggt gat gcc gct			864
Ala Glu Pro Met Arg Phe Gly Arg Leu Phe Leu Ala Gly Asp Ala Ala			
275	280	285	
cac att gtt ccg cct acc gga gca aag ggc ttg aat ctg gcg gct tca			912
His Ile Val Pro Pro Thr Gly Ala Lys Gly Leu Asn Leu Ala Ala Ser			
290	295	300	
gac gtg tac tac ctg agc cgt gca atg cgg gag tac tat gcc gag aag			960
Asp Val Tyr Tyr Leu Ser Arg Ala Met Arg Glu Tyr Tyr Ala Glu Lys			
305	310	315	320
tcc gaa gcc ggt atc gac gct tac tcg gca tcc gcc ttg cgc cgt gtt			1008
Ser Glu Ala Gly Ile Asp Ala Tyr Ser Ala Ser Ala Leu Arg Arg Val			
325	330	335	
tgg aag gcg gaa cgt ttc tcc tgg tgg atg act tct cag ctg cac cgc			1056
Trp Lys Ala Glu Arg Phe Ser Trp Trp Met Thr Ser Gln Leu His Arg			
340	345	350	
ttt cca gat tcc gac gca ttc agc caa cgc atc cag act gct gag ctc			1104
Phe Pro Asp Ser Asp Ala Phe Ser Gln Arg Ile Gln Thr Ala Glu Leu			
355	360	365	
gat tat ctg gtg aac tcc aaa gca gcc att acc tcc ttg gcc gag aac			1152
Asp Tyr Leu Val Asn Ser Lys Ala Ala Ile Thr Ser Leu Ala Glu Asn			
370	375	380	
tac gta ggc ctg cca tac taa			1173
Tyr Val Gly Leu Pro Tyr			
385	390		
<210> 6			

<211> 390

<212> PRT

<213> Rhodopseudomonas palustris

<400> 6

Met Arg Thr Gln Val Ala Ile Ile Gly Ala Gly Pro Ser Gly Leu Leu

1 5 10 15

Leu Gly Gln Leu Leu His Lys Tyr Gly Ile Asp Ala Val Ile Val Glu

20 25 30

Arg Lys Asp Pro Asp Tyr Val Leu Ser Arg Ile Arg Ala Gly Val Leu

35 40 45

Glu Gln Gly Met Val Asp Leu Leu Asp Glu Ala Gly Val Ser Ala Arg

50 55 60

Leu His Gln Glu Ala Leu Val His Gly Gly Phe Glu Ile Ala Phe Ala

65 70 75 80

Gly Gln Arg His Pro Ile Asp Leu Arg Gly Ala Thr Gly Gly Lys Ser

85 90 95

Val Thr Val Tyr Gly Gln Thr Glu Val Thr Arg Asp Leu Met Glu Ala

100 105 110

Arg Ser Ala Ala Gly Leu Thr Thr Ile Tyr Asp Ala Ala Asp Val Ser

115 120 125

Leu His Asp Phe Glu Gly Ala His Pro Lys Val Arg Tyr Val Lys Asp

130 135 140

Gly Thr Thr Arg Glu Ile Val Cys Asp Phe Ile Ala Gly Cys Asp Gly

145 150 155 160

Phe His Gly Ile Ser Arg Gln Ser Val Pro Ala Ser Ala Val Gln Ser

165 170 175

Phe Glu Arg Val Tyr Pro Phe Gly Trp Leu Gly Leu Leu Ser Asp Thr

180 185 190

Pro Pro Val Ser Pro Glu Leu Ile Tyr Val Asn His Asp Arg Gly Phe

195 200 205
 Ala Leu Cys Ser Met Arg Ser Met His Arg Ser Arg Tyr Tyr Val Gln
 210 215 220
 Cys Pro Leu Thr Asp Asp Val Ala Asp Trp Ser Asp Asp Arg Phe Trp
 225 230 235 240
 Asp Glu Leu Lys Ser Arg Leu Asp Pro Glu Thr Ala Gly Lys Leu Val
 245 250 255
 Thr Gly Pro Ser Ile Glu Lys Ser Ile Ala Pro Leu Arg Ser Phe Val
 260 265 270
 Ala Glu Pro Met Arg Phe Gly Arg Leu Phe Leu Ala Gly Asp Ala Ala
 275 280 285
 His Ile Val Pro Pro Thr Gly Ala Lys Gly Leu Asn Leu Ala Ala Ser
 290 295 300
 Asp Val Tyr Tyr Leu Ser Arg Ala Met Arg Glu Tyr Tyr Ala Glu Lys
 305 310 315 320
 Ser Glu Ala Gly Ile Asp Ala Tyr Ser Ala Ser Ala Leu Arg Arg Val
 325 330 335
 Trp Lys Ala Glu Arg Phe Ser Trp Trp Met Thr Ser Gln Leu His Arg
 340 345 350
 Phe Pro Asp Ser Asp Ala Phe Ser Gln Arg Ile Gln Thr Ala Glu Leu
 355 360 365
 Asp Tyr Leu Val Asn Ser Lys Ala Ala Ile Thr Ser Leu Ala Glu Asn
 370 375 380
 Tyr Val Gly Leu Pro Tyr
 385 390
 <210> 7
 <211> 1179
 <212> ADN
 <213> Bradyrhizobium diazoefficiens

<220>

<221> CDS

<222> (1)..(1176)

<223> Oligonucleotit được tối ưu hóa cho codon

<400> 7

```

atg cgt act cag gtg gga atc gtg gga gcc gga cca gcc ggt ctg ctc      48
Met Arg Thr Gln Val Gly Ile Val Gly Ala Gly Pro Ala Gly Leu Leu
1           5           10          15
ctc tcg cac atg ctg tat ctg tcc ggc att gag tcg atc atc atc gag      96
Leu Ser His Met Leu Tyr Leu Ser Gly Ile Glu Ser Ile Ile Ile Glu
           20          25          30
tca cgt tcg cag gcg gaa atc gag cag acc atc cgc gcc ggg gta ctg     144
Ser Arg Ser Gln Ala Glu Ile Glu Gln Thr Ile Arg Ala Gly Val Leu
          35          40          45
gaa cag tct acg gtt gac ctg atg acc gag att ggg gct ggc gat cgc     192
Glu Gln Ser Thr Val Asp Leu Met Thr Glu Ile Gly Ala Gly Asp Arg
          50          55          60
atg aag cgg gaa ggc ttt gtc cat ggt ggc ttt gag ctc cgc ttt gca     240
Met Lys Arg Glu Gly Phe Val His Gly Gly Phe Glu Leu Arg Phe Ala
          65          70          75          80
ggg cat ggg cac cgg atc gat ctg cag aac ctt gct aat ggc cgc acg     288
Gly His Gly His Arg Ile Asp Leu Gln Asn Leu Ala Asn Gly Arg Thr
          85          90          95
atc aca gtc tat cca cag cac gaa gtc ctc aag gac ctg att gcc ttg     336
Ile Thr Val Tyr Pro Gln His Glu Val Leu Lys Asp Leu Ile Ala Leu
          100         105         110
cgc ttg cgt act ggt ggt caa atc cac ttt gag gct aaa gct acc tcc     384
Arg Leu Arg Thr Gly Gly Gln Ile His Phe Glu Ala Lys Ala Thr Ser
          115         120         125

```

att gac ggt ctg acc agc gat cag cca gtt gtg cgc ttc att acc aag 432
 Ile Asp Gly Leu Thr Ser Asp Gln Pro Val Val Arg Phe Ile Thr Lys
 130 135 140
 caa ggt gaa acc cgt gaa ttg tcc tgc gac ttt gtg gct ggc tgt gac 480
 Gln Gly Glu Thr Arg Glu Leu Ser Cys Asp Phe Val Ala Gly Cys Asp
 145 150 155 160
 ggt ggc tat gga gcc tca cga gct gcg att ccg gag gat ctg gtt cgc 528
 Gly Gly Tyr Gly Ala Ser Arg Ala Ala Ile Pro Glu Asp Leu Val Arg
 165 170 175
 cgc gat tac ttc cga gtg tac ccg ttc ggt tgg ttc ggc atc ctg gca 576
 Arg Asp Tyr Phe Arg Val Tyr Pro Phe Gly Trp Phe Gly Ile Leu Ala
 180 185 190
 aaa gcc cca ccc tca tcc gaa gag ctc atc tac gca cat cac gac cgc 624
 Lys Ala Pro Pro Ser Ser Glu Glu Leu Ile Tyr Ala His His Asp Arg
 195 200 205
 gga ttc gcg ctc atc agc aca cgc tct ccg aat atc cag cgc atg tac 672
 Gly Phe Ala Leu Ile Ser Thr Arg Ser Pro Asn Ile Gln Arg Met Tyr
 210 215 220
 ttc cag tgc gca cct acc gat tcc gtt gac aac tgg agc gat gat cgc 720
 Phe Gln Cys Ala Pro Thr Asp Ser Val Asp Asn Trp Ser Asp Asp Arg
 225 230 235 240
 att tgg aac gag ttg cag act cgg gtt ggt ggc gat ggc ttc gag ctc 768
 Ile Trp Asn Glu Leu Gln Thr Arg Val Gly Gly Asp Gly Phe Glu Leu
 245 250 255
 aag act ggc cct atc ttc cag aag ggt atc att cct ctt cgc tcc ttc 816
 Lys Thr Gly Pro Ile Phe Gln Lys Gly Ile Ile Pro Leu Arg Ser Phe
 260 265 270
 gtt tgc gaa ccc atg caa cac gga cgg ctt ttc ctt gca ggc gat gct 864
 Val Cys Glu Pro Met Gln His Gly Arg Leu Phe Leu Ala Gly Asp Ala


```

      275              280              285
gca cac tct gtc cct ccc aca ggt gca aaa ggc ctc aac ctg gca gct      912
Ala His Ser Val Pro Pro Thr Gly Ala Lys Gly Leu Asn Leu Ala Ala

      290              295              300
gca gac gtg tac gtc ctt gcc cgt gca ctt gcg agc tac tac gct aag      960
Ala Asp Val Tyr Val Leu Ala Arg Ala Leu Ala Ser Tyr Tyr Ala Lys

      305              310              315              320
cga tcc acc acc ctg ttg gac gcg tat tcc tct acc gca ttg cgt cgc      1008
Arg Ser Thr Thr Leu Leu Asp Ala Tyr Ser Ser Thr Ala Leu Arg Arg

              325              330              335
gtg tgg cgt gcc caa cac ttc tcc tgg tgg atg acc tct atg ctg cat      1056
Val Trp Arg Ala Gln His Phe Ser Trp Trp Met Thr Ser Met Leu His

              340              345              350
tgg ttt cac gaa gga acc gaa ttc gac ctg aag cgt caa ctg gcg gag      1104
Trp Phe His Glu Gly Thr Glu Phe Asp Leu Lys Arg Gln Leu Ala Glu

              355              360              365
ctc gaa ttg gtc acg tcc tcc aaa gcg gct gca acc act ctg gcc gaa      1152
Leu Glu Leu Val Thr Ser Ser Lys Ala Ala Ala Thr Thr Leu Ala Glu

      370              375              380
aac tac gta ggc atg cca ctt gct taa      1179
Asn Tyr Val Gly Met Pro Leu Ala

      385              390
<210> 8
<211> 392
<212> PRT
<213> Bradyrhizobium diazoefficiens
<400> 8
Met Arg Thr Gln Val Gly Ile Val Gly Ala Gly Pro Ala Gly Leu Leu

```

1

5

10

15

Leu Ser His Met Leu Tyr Leu Ser Gly Ile Glu Ser Ile Ile Ile Glu
 20 25 30
 Ser Arg Ser Gln Ala Glu Ile Glu Gln Thr Ile Arg Ala Gly Val Leu
 35 40 45
 Glu Gln Ser Thr Val Asp Leu Met Thr Glu Ile Gly Ala Gly Asp Arg
 50 55 60
 Met Lys Arg Glu Gly Phe Val His Gly Gly Phe Glu Leu Arg Phe Ala
 65 70 75 80
 Gly His Gly His Arg Ile Asp Leu Gln Asn Leu Ala Asn Gly Arg Thr
 85 90 95
 Ile Thr Val Tyr Pro Gln His Glu Val Leu Lys Asp Leu Ile Ala Leu
 100 105 110
 Arg Leu Arg Thr Gly Gly Gln Ile His Phe Glu Ala Lys Ala Thr Ser
 115 120 125
 Ile Asp Gly Leu Thr Ser Asp Gln Pro Val Val Arg Phe Ile Thr Lys
 130 135 140
 Gln Gly Glu Thr Arg Glu Leu Ser Cys Asp Phe Val Ala Gly Cys Asp
 145 150 155 160
 Gly Gly Tyr Gly Ala Ser Arg Ala Ala Ile Pro Glu Asp Leu Val Arg
 165 170 175
 Arg Asp Tyr Phe Arg Val Tyr Pro Phe Gly Trp Phe Gly Ile Leu Ala
 180 185 190
 Lys Ala Pro Pro Ser Ser Glu Glu Leu Ile Tyr Ala His His Asp Arg
 195 200 205
 Gly Phe Ala Leu Ile Ser Thr Arg Ser Pro Asn Ile Gln Arg Met Tyr
 210 215 220
 Phe Gln Cys Ala Pro Thr Asp Ser Val Asp Asn Trp Ser Asp Asp Arg
 225 230 235 240
 Ile Trp Asn Glu Leu Gln Thr Arg Val Gly Gly Asp Gly Phe Glu Leu

```

                245                250                255
Lys Thr Gly Pro Ile Phe Gln Lys Gly Ile Ile Pro Leu Arg Ser Phe

                260                265                270
Val Cys Glu Pro Met Gln His Gly Arg Leu Phe Leu Ala Gly Asp Ala

                275                280                285
Ala His Ser Val Pro Pro Thr Gly Ala Lys Gly Leu Asn Leu Ala Ala

                290                295                300
Ala Asp Val Tyr Val Leu Ala Arg Ala Leu Ala Ser Tyr Tyr Ala Lys

305                310                315                320
Arg Ser Thr Thr Leu Leu Asp Ala Tyr Ser Ser Thr Ala Leu Arg Arg

                325                330                335
Val Trp Arg Ala Gln His Phe Ser Trp Trp Met Thr Ser Met Leu His

                340                345                350
Trp Phe His Glu Gly Thr Glu Phe Asp Leu Lys Arg Gln Leu Ala Glu

                355                360                365
Leu Glu Leu Val Thr Ser Ser Lys Ala Ala Ala Thr Thr Leu Ala Glu

                370                375                380
Asn Tyr Val Gly Met Pro Leu Ala

385                390

<210> 9
<211> 1179
<212> ADN
<213> Cupriavidus metallidurans
<220>
<221> CDS
<222> (1)..(1176)
<223> Oligonucleotit được tối ưu hóa cho codon
<400> 9

atg cgc act cag gtt ggt atc att gga gct ggt cca gca ggc ttg ctc

```

Met Arg Thr Gln Val Gly Ile Ile Gly Ala Gly Pro Ala Gly Leu Leu
1 5 10 15
ctt tcc cat ctc ctg cac ctc aag ggt atc gat tct gtc gtc atc gaa 96
Leu Ser His Leu Leu His Leu Lys Gly Ile Asp Ser Val Val Ile Glu
20 25 30
tct cgg aca cgc gaa gag atc gag tcc acg att cgg gcg ggt gta ctg 144
Ser Arg Thr Arg Glu Glu Ile Glu Ser Thr Ile Arg Ala Gly Val Leu
35 40 45
gaa cag ggc acc atg gac ctc ttg cag gat gtt ggc ctt ggg aat cgc 192
Glu Gln Gly Thr Met Asp Leu Leu Gln Asp Val Gly Leu Gly Asn Arg
50 55 60
atg gtc gcc gag ggc gca ctt cac caa ggc ttc gaa ttg acc ttt gag 240
Met Val Ala Glu Gly Ala Leu His Gln Gly Phe Glu Leu Thr Phe Glu
65 70 75 80
ggc gaa cgg cat cgt atc gat ctc acc gac ctg acg ggc cat gcg atc 288
Gly Glu Arg His Arg Ile Asp Leu Thr Asp Leu Thr Gly His Ala Ile
85 90 95
acc gta tac gca cag cac gaa gtg atc aag gat ctt gtg gca gca cga 336
Thr Val Tyr Ala Gln His Glu Val Ile Lys Asp Leu Val Ala Ala Arg
100 105 110
gtg gca gct ggt ggg aaa ctg gtt ttc ggc gtt tcg aac act tct ctc 384
Val Ala Ala Gly Gly Lys Leu Val Phe Gly Val Ser Asn Thr Ser Leu
115 120 125
cac gat gtg gaa tcc acc aaa ccg tcg att cgc tac gag aag gac gac 432
His Asp Val Glu Ser Thr Lys Pro Ser Ile Arg Tyr Glu Lys Asp Asp
130 135 140
gac acg tgc gaa att cac tgc gac ttc gtc atc ggg tgt gat ggc tca 480
Asp Thr Cys Glu Ile His Cys Asp Phe Val Ile Gly Cys Asp Gly Ser
145 150 155 160

```

caa gga cca tca cgc tct gca att ccc gaa gca gtc cgc aaa gac cat      528
Gln Gly Pro Ser Arg Ser Ala Ile Pro Glu Ala Val Arg Lys Asp His
      165              170              175
cag cgc ttg tac ccc ttt ggc tgg ttt ggc att ctg gtc gaa gct ccg      576
Gln Arg Leu Tyr Pro Phe Gly Trp Phe Gly Ile Leu Val Glu Ala Pro
      180              185              190
cct tca tcc gac gag ctg atc tac gca cga cat gac cgc gga ttt gcc      624
Pro Ser Ser Asp Glu Leu Ile Tyr Ala Arg His Asp Arg Gly Phe Ala
      195              200              205
ctg atc tcg act cgt agc cct acc gtc caa cgt atg tac ttc cag tgc      672
Leu Ile Ser Thr Arg Ser Pro Thr Val Gln Arg Met Tyr Phe Gln Cys
      210              215              220
gat cca cgg gat agc gtt gag aac tgg tcc gac gat cgc atc tgg tcc      720
Asp Pro Arg Asp Ser Val Glu Asn Trp Ser Asp Asp Arg Ile Trp Ser
      225              230              235              240
gag ctc cac gca cga ttg gat caa gcc gat ggt tgg cgt gta act gag      768
Glu Leu His Ala Arg Leu Asp Gln Ala Asp Gly Trp Arg Val Thr Glu
      245              250              255
ggc cgc att ttc cag aag aac atc gtt ggt atg cgc tcc ttc gtg agc      816
Gly Arg Ile Phe Gln Lys Asn Ile Val Gly Met Arg Ser Phe Val Ser
      260              265              270
aat gtg atg cag cac ggt cgc ctg ttt ctg gct ggg gat tct gcc cac      864
Asn Val Met Gln His Gly Arg Leu Phe Leu Ala Gly Asp Ser Ala His
      275              280              285
atc gtt ccg cct act ggc gcg aaa gga atg aac ctc gcc gtg aac gat      912
Ile Val Pro Pro Thr Gly Ala Lys Gly Met Asn Leu Ala Val Asn Asp
      290              295              300
gtc aag atc ctg gct gaa ggt ctg gac tcc ttc tac aag aac ggt acc      960
Val Lys Ile Leu Ala Glu Gly Leu Asp Ser Phe Tyr Lys Asn Gly Thr

```

305 310 315 320
 gag gac aag ctg aat gcg tat acc gcc acc gcc ctg cag cgt atc tgg 1008
 Glu Asp Lys Leu Asn Ala Tyr Thr Ala Thr Ala Leu Gln Arg Ile Trp
 325 330 335
 cgt gcg gag cac ttc tcc tgg tgg atg acc tcc atg ttg cac cgc ttc 1056
 Arg Ala Glu His Phe Ser Trp Trp Met Thr Ser Met Leu His Arg Phe
 340 345 350
 gct gat gcg acc cca ttc gac cag caa ctt cag gtg tcc gaa ctg cgc 1104
 Ala Asp Ala Thr Pro Phe Asp Gln Gln Leu Gln Val Ser Glu Leu Arg
 355 360 365
 tat gtc acc tcg tcc cgt gct ggc gct aca gct ctt gcc gag aac tat 1152
 Tyr Val Thr Ser Ser Arg Ala Gly Ala Thr Ala Leu Ala Glu Asn Tyr
 370 375 380
 gtg gga atg gtt gga ctg agc cac taa 1179
 Val Gly Met Val Gly Leu Ser His
 385 390
 <210> 10
 <211> 392
 <212> PRT
 <213> *Cupriavidus metallidurans*
 <400> 10
 Met Arg Thr Gln Val Gly Ile Ile Gly Ala Gly Pro Ala Gly Leu Leu
 1 5 10 15
 Leu Ser His Leu Leu His Leu Lys Gly Ile Asp Ser Val Val Ile Glu
 20 25 30
 Ser Arg Thr Arg Glu Glu Ile Glu Ser Thr Ile Arg Ala Gly Val Leu
 35 40 45
 Glu Gln Gly Thr Met Asp Leu Leu Gln Asp Val Gly Leu Gly Asn Arg
 50 55 60

Met Val Ala Glu Gly Ala Leu His Gln Gly Phe Glu Leu Thr Phe Glu
 65 70 75 80
 Gly Glu Arg His Arg Ile Asp Leu Thr Asp Leu Thr Gly His Ala Ile
 85 90 95
 Thr Val Tyr Ala Gln His Glu Val Ile Lys Asp Leu Val Ala Ala Arg
 100 105 110
 Val Ala Ala Gly Gly Lys Leu Val Phe Gly Val Ser Asn Thr Ser Leu
 115 120 125
 His Asp Val Glu Ser Thr Lys Pro Ser Ile Arg Tyr Glu Lys Asp Asp
 130 135 140
 Asp Thr Cys Glu Ile His Cys Asp Phe Val Ile Gly Cys Asp Gly Ser
 145 150 155 160
 Gln Gly Pro Ser Arg Ser Ala Ile Pro Glu Ala Val Arg Lys Asp His
 165 170 175
 Gln Arg Leu Tyr Pro Phe Gly Trp Phe Gly Ile Leu Val Glu Ala Pro
 180 185 190
 Pro Ser Ser Asp Glu Leu Ile Tyr Ala Arg His Asp Arg Gly Phe Ala
 195 200 205
 Leu Ile Ser Thr Arg Ser Pro Thr Val Gln Arg Met Tyr Phe Gln Cys
 210 215 220
 Asp Pro Arg Asp Ser Val Glu Asn Trp Ser Asp Asp Arg Ile Trp Ser
 225 230 235 240
 Glu Leu His Ala Arg Leu Asp Gln Ala Asp Gly Trp Arg Val Thr Glu
 245 250 255
 Gly Arg Ile Phe Gln Lys Asn Ile Val Gly Met Arg Ser Phe Val Ser
 260 265 270
 Asn Val Met Gln His Gly Arg Leu Phe Leu Ala Gly Asp Ser Ala His
 275 280 285
 Ile Val Pro Pro Thr Gly Ala Lys Gly Met Asn Leu Ala Val Asn Asp

290 295 300
 Val Lys Ile Leu Ala Glu Gly Leu Asp Ser Phe Tyr Lys Asn Gly Thr
 305 310 315 320
 Glu Asp Lys Leu Asn Ala Tyr Thr Ala Thr Ala Leu Gln Arg Ile Trp
 325 330 335
 Arg Ala Glu His Phe Ser Trp Trp Met Thr Ser Met Leu His Arg Phe
 340 345 350
 Ala Asp Ala Thr Pro Phe Asp Gln Gln Leu Gln Val Ser Glu Leu Arg
 355 360 365
 Tyr Val Thr Ser Ser Arg Ala Gly Ala Thr Ala Leu Ala Glu Asn Tyr
 370 375 380
 Val Gly Met Val Gly Leu Ser His
 385 390
 <210> 11
 <211> 1176
 <212> ADN
 <213> Rhodococcus sp.
 <220>
 <221> CDS
 <222> (1)..(1173)
 <223> Oligonucleotit được tối ưu hóa cho codon
 <400> 11
 atg cgt acc caa gtg gcc atc att gga gcg ggt cca gct ggg ctg ctg 48
 Met Arg Thr Gln Val Ala Ile Ile Gly Ala Gly Pro Ala Gly Leu Leu
 1 5 10 15
 ctc agc cac ctc ctg gat gaa cag gga atc gac tca atc ctg atc gaa 96
 Leu Ser His Leu Leu Asp Glu Gln Gly Ile Asp Ser Ile Leu Ile Glu
 20 25 30
 tct cgc act cag gaa tac gtt ctg tca cgc atc cgt gcc ggt gtc ctg 144

Ser Arg Thr Gln Glu Tyr Val Leu Ser Arg Ile Arg Ala Gly Val Leu
 35 40 45
 gag cac tcc acg gtg caa ctt ctg gat gag cat ggg ttg ggt gag cgc 192
 Glu His Ser Thr Val Gln Leu Leu Asp Glu His Gly Leu Gly Glu Arg
 50 55 60
 ctg cat cgc gaa ggt gat gaa cat cgc ggc atc tac ttg cag tgg ccc 240
 Leu His Arg Glu Gly Asp Glu His Arg Gly Ile Tyr Leu Gln Trp Pro
 65 70 75 80
 gaa gaa cga cac cac atc gac ttc cgg gac ctg gtc gat cgt tcc gtc 288
 Glu Glu Arg His His Ile Asp Phe Arg Asp Leu Val Asp Arg Ser Val
 85 90 95
 tgg gtg tat ggt cag acc gag gtg aca aag gat ctg gtc gtc gca cgc 336
 Trp Val Tyr Gly Gln Thr Glu Val Thr Lys Asp Leu Val Val Ala Arg
 100 105 110
 gag aaa gcg ggt caa cag atc tac tac gat gtg tcc gac acc gcg ctt 384
 Glu Lys Ala Gly Gln Gln Ile Tyr Tyr Asp Val Ser Asp Thr Ala Leu
 115 120 125
 cac gac gta gaa tcc gac gca ccc tac gtt acc ttc act gac gca tcg 432
 His Asp Val Glu Ser Asp Ala Pro Tyr Val Thr Phe Thr Asp Ala Ser
 130 135 140
 ggc aat gcg gtt cgc att gat gca acc gtt gtt gcg ggc tgt gat ggc 480
 Gly Asn Ala Val Arg Ile Asp Ala Thr Val Val Ala Gly Cys Asp Gly
 145 150 155 160
 tct ttc ggt cca tca cgg gct gca atg cct gac tcg gtt cgt aac acc 528
 Ser Phe Gly Pro Ser Arg Ala Ala Met Pro Asp Ser Val Arg Asn Thr
 165 170 175
 tgg gag cgt gtg tac cca tat tcc tgg ttg ggc gtg ctt gca gat gtg 576
 Trp Glu Arg Val Tyr Pro Tyr Ser Trp Leu Gly Val Leu Ala Asp Val
 180 185 190

gct cct tct acc gac gag ctg atc tat gcc tgg cat cag gac ggt ttt 624
 Ala Pro Ser Thr Asp Glu Leu Ile Tyr Ala Trp His Gln Asp Gly Phe
 195 200 205
 gca atg cac tcc atg cga tcc tcg acc gtt tct cgc ctg tac ctc cag 672
 Ala Met His Ser Met Arg Ser Ser Thr Val Ser Arg Leu Tyr Leu Gln
 210 215 220
 gtt cct aac ggg act gac att gac acc tgg tcc gac gac cgc atc tgg 720
 Val Pro Asn Gly Thr Asp Ile Asp Thr Trp Ser Asp Asp Arg Ile Trp
 225 230 235 240
 gat gct ctg gcc ctc cgt ctt gga cac gga caa gat ggc tgg acc ctg 768
 Asp Ala Leu Ala Leu Arg Leu Gly His Gly Gln Asp Gly Trp Thr Leu
 245 250 255
 aat ccc ggc ccg att acc gag aag tcg gtg ttg cca atg cgc tct tac 816
 Asn Pro Gly Pro Ile Thr Glu Lys Ser Val Leu Pro Met Arg Ser Tyr
 260 265 270
 gtc cag act cca atg cgc cat ggc aac ctt tat ctg gct ggt gat gca 864
 Val Gln Thr Pro Met Arg His Gly Asn Leu Tyr Leu Ala Gly Asp Ala
 275 280 285
 gct cac atc gtc ccg cct act ggc gct aag ggt ctg aac ctg gct gta 912
 Ala His Ile Val Pro Pro Thr Gly Ala Lys Gly Leu Asn Leu Ala Val
 290 295 300
 gca gat gtc gca ctc ctc gca cca gcc ttg gcg caa aag ctc aaa ggc 960
 Ala Asp Val Ala Leu Leu Ala Pro Ala Leu Ala Gln Lys Leu Lys Gly
 305 310 315 320
 aac gac tcc cgt gcc gcg gat agc tac agc gat gat gcc ttg cga cgg 1008
 Asn Asp Ser Arg Ala Ala Asp Ser Tyr Ser Asp Asp Ala Leu Arg Arg
 325 330 335
 gta tgg cgc tgc acc cac ttc agc tgg tgg atg acg acg atg ctt cac 1056
 Val Trp Arg Cys Thr His Phe Ser Trp Trp Met Thr Thr Met Leu His

```

          340          345          350
aca gga gat gac ccg ttt gat gcc cag ctc cag ctt tcc cag ctc aag      1104
Thr Gly Asp Asp Pro Phe Asp Ala Gln Leu Gln Leu Ser Gln Leu Lys

          355          360          365
tgg gtc gca tcc tcc gaa gcc gga gct atg ggc ttg gct gag aac tac      1152
Trp Val Ala Ser Ser Glu Ala Gly Ala Met Gly Leu Ala Glu Asn Tyr

          370          375          380
gct ggc ctt ccg att ggc ttc taa      1176
Ala Gly Leu Pro Ile Gly Phe

385          390
<210> 12
<211> 391
<212> PRT
<213> Rhodococcus sp.
<400> 12
Met Arg Thr Gln Val Ala Ile Ile Gly Ala Gly Pro Ala Gly Leu Leu
1          5          10          15
Leu Ser His Leu Leu Asp Glu Gln Gly Ile Asp Ser Ile Leu Ile Glu
          20          25          30
Ser Arg Thr Gln Glu Tyr Val Leu Ser Arg Ile Arg Ala Gly Val Leu
          35          40          45
Glu His Ser Thr Val Gln Leu Leu Asp Glu His Gly Leu Gly Glu Arg
          50          55          60
Leu His Arg Glu Gly Asp Glu His Arg Gly Ile Tyr Leu Gln Trp Pro
65          70          75          80
Glu Glu Arg His His Ile Asp Phe Arg Asp Leu Val Asp Arg Ser Val
          85          90          95
Trp Val Tyr Gly Gln Thr Glu Val Thr Lys Asp Leu Val Val Ala Arg
          100          105          110

```

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

340 345 350
 Thr Gly Asp Asp Pro Phe Asp Ala Gln Leu Gln Leu Ser Gln Leu Lys
 355 360 365
 Trp Val Ala Ser Ser Glu Ala Gly Ala Met Gly Leu Ala Glu Asn Tyr
 370 375 380
 Ala Gly Leu Pro Ile Gly Phe

385 390

<210> 13

<211> 42

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 13

tattaattaa atgcgcgttt taattattga taattatgat tc

42

<210> 14

<211> 45

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 14

ttgcggcgcc ttgttttaaac ctccttacag aaaaatgggtt gggcg

45

<210> 15

<211> 33

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 15

aaattttaaac ctcctttaca gaaaaatggg tgg

33

<210> 16

<211> 28

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 16

ggaggttttaa acaagcggcc gcgatatc

28

<210> 17

<211> 35

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 17

aggaggttta aatttatgag cactcaggtg gctat

35

<210> 18

<211> 35

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 18

cttggtttaa cctccttata cgagtggcag tccta

35

<210> 19

<211> 31

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 19

gctatcaaaa cattcggcac attggttttc c

31

<210> 20

<211> 29

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 20

ggaagatgcg tgatctgac cttcaactc

29

<210> 21

<211> 24

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 21

ttacagaaaa atggttgggc gcaa

24

<210> 22

<211> 22

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 22

atgcgcactc aggtggctat cg

22

<210> 23

<211> 342

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 23

tacgtacctg caggtacgt gtcagtaggc gcgtagggta agtggggtag cggcttggtta	60
gatatcttga aatcggcttt caacagcatt gatttcgatg tatttagctg gccgttacct	120
tgcgaaatgtc cacagggtag ctggtagttt gaaaatcaac gccgttgccc ttaggattca	180
gtaactggca cattttgtaa tgcgctagat ctgtgtgctc agtcttccag gctgcttata	240
acagtgaaag caaaaccaat tcgtggctgc gaaagtcgta gccaccacga agtccaaagg	300
aggatctaaa ttatgaataa tataaaagga ggaattaatt aa	342

<210> 24

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 24

accatttttc tgtaatacgt acctgcaggt agcgtg	36
---	----

<210> 25

<211> 35

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 25

cacctgagtg cgcatttaata taattcctcc tttta	35
---	----

<210> 26

<211> 29

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 26

gcttgtaga tatcttgaaa tcggctttc

29

<210> 27

<211> 18

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 27

ggaggttttaa acaagcgg

18

<210> 28

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 28

aatttagatc ctcctttgga cttcgtg

27

<210> 29

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 29

aggaggatct aaattatgcg tactcaggtg ggaatc

36

<210> 30

<211> 32

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 30

cttggttaaaa cctccttaag caagtggcat gc

32

<210> 31

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 31

aggaggatct aaattatgcg cactcaggtg gcaatc

36

<210> 32

<211> 34

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 32

cttggttaaaa cctccttagt atggcaggcc tacg

34

<210> 33

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 33

aggaggatct aaattatgcg cacccaagtg gtcac

36

<210> 34

<211> 37

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 34

cttggtttaa cctccttaga acggcagacc cacgtag

37

<210> 35

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 35

aggaggatct aaattatgcg cactcaggtt ggtatc

36

<210> 36

<211> 38

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 36

cttggtttaaa cctccttagt ggctcagtc aaccattc

38

<210> 37

<211> 40

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 37

aggaggatct aaattatgcg tacccaagtg gccatcattg

40

<210> 38

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 38

cttggtttaaa cctccttaga agccaatcgg aaggcc

36

<210> 39

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 39

gccgggctcc tggaacagtc tacggtt

27

<210> 40

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 40

ttccaggagc ccggcgcgga tggctctg

27

<210> 41

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 41

gctggtctcc tggaacgtat cacggtg

27

<210> 42

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 42

ttccaggaga ccagcccgaa ctcggcc

27

<210> 43

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 43

gcaggcctcc tggagcaggg catggtt

27

<210> 44

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 44

ctccaggagg cctgcacgga tgcggga

27

<210> 45

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 45

gctggactct tggaacaggg caccgtt

27

<210> 46

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 46

ttccaagagt ccagcgcgaa ctgccc

27

<210> 47

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 47

gcgggtctcc tggaacaggg caccatg

27

<210> 48

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 48

ttccaggaga ccgcccga tctgtga

27

<210> 49

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 49

gcgggtctcc tggagcactc cacggtg

27

<210> 50

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 50

ctccaggaga ccggcacgga tgcgtga

27

<210> 51

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 51

ctcatcttcg cacatcacga ccgcgga

27

<210> 52

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 52

atgtgcgaag atgagctctt cggatga

27

<210> 53

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 53

cgcatgttct tccagtgcgc acctacc

27

<210> 54

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 54

ctggaagaac atgcgctgga tattcgg

27

<210> 55

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 55

ttgatcttct cgaaccatga tcgcggt

27

<210> 56

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 56

gttcgagaag atcaactcgt ggtcaca

27

<210> 57

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 57

cgctatttcg tgcagtgctc actcgac

27

<210> 58

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 58

ctgcacgaaa tagcgggagc gtgtcgg

27

<210> 59

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 59

ctgatcttcg tcaaccacga ccgaggc

27

<210> 60

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 60

gttgacgaag atcagttctg ggctgac

27

<210> 61

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 61

cggtacttcg tccaatgcc tttgacc

27

<210> 62

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 62

ttggacgaag taccgtagaac ggtgcat

27

<210> 63

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 63

ctcgtgttcg ctaatcaccc acgcggg

27

<210> 64

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 64

attagcgaac acgagttcat gatcgac

27

<210> 65

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 65

cgctacttca tccagtgcc tttggag

27

<210> 66

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 66

ctggatgaag tagcgagaac gggatatg

27

<210> 67

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 67

gctggtattc tggaacgtat cacggtg

27

<210> 68

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 68

ttccagaata ccagcccgaa ctcggcc

27

<210> 69

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 69

gctggttctc tggaacgtat cacggtg

27

<210> 70

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 70

ttccagagaa ccagcccgaa ctcggcc

27

<210> 71

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 71

gctggtacac tggaacgtat cacggtg

27

<210> 72

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 72

ttccagtgtgta ccagcccgaa ctcggcc

27

<210> 73

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 73

gctggttgctc tggaacgtat cacggtg

27

<210> 74

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 74

ttccagacaa ccagcccgaa ctcggcc

27

<210> 75

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 75

gctggtatgc tggaacgtat cacggtg

27

<210> 76

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 76

ttccagcata ccagcccgaa ctcggcc

27

<210> 77

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 77

gctggtcaac tggaacgtat cacggtg

27

<210> 78

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 78

ttccagttga ccagcccgaa ctcggcc

27

<210> 79

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 79

ttggtggcag ctggcgctaa tcttgcg

27

<210> 80

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 80

gccagctgcc accaagccct ctcggcg

27

<210> 81

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 81

ttggtgatgg ctggcgctaa tcttgcg

27

<210> 82

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 82

gccagccatc accaagccct ctcggcg

27

<210> 83

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 83

tttgccatgg cgtcgatgcg ctcaccg

27

<210> 84

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 84

cgacgccatg gcaaaaccgc gatcatg

27

<210> 85

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 85

ccacccgcag gagcgaaagg gatgaac

27

<210> 86

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 86

cgctcctgcg ggtggaacga tatgagc

27

<210> 87

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 87

ccacccggtg gagcgaaagg gatgaac

27

<210> 88

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 88

cgctccaccg ggtggaacga tatgagc

27

<210> 89

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 89

ccaccctgtg gagcgaaagg gatgaac

27

<210> 90

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 90

cgctccacag ggtggaacga tatgagc

27

<210> 91

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 91

ccaccctctg gagcgaaagg gatgaac

27

<210> 92

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 92

cgctccagag ggtggaacga tatgagc

27

<210> 93

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 93

gagaacgttg taggactgcc actcgta

27

<210> 94

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 94

tectacaacg ttctccgcca gggtagc

27

<210> 95

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 95

gagaacctcg taggactgcc actcgta

27

<210> 96

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 96

tctacgagg ttctccgcca gggtagc

27

<210> 97

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 97

gagaacattg taggactgcc actcgta

27

<210> 98

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 98

tcctacaatg ttctccgcca gggtgac

27

<210> 99

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 99

gagaacatgg taggactgcc actcgta

27

<210> 100

<211> 27

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Oligonucleotit tổng hợp

<400> 100

tcctaccatg ttctccgcca gggtgac

27